DOCUMENT RESUME

ED 232 869 SE 042 643

AUTHOR	Blosser, Patricia E., Ed.; Mayer, Victor J., Ed.
TITLE	Investigations in Science Education. Volume 9, Number 3.
INSTITUTION	ERIC Clearinghouse for Science, Mathematics, and Environmental Education, Columbus, Ohio.; Ohio State Univ., Columbus. Center for Science and Mathematics Education.
SPONS AGENCY	National Inst. of Education (ED), Washington, DC.
PUB DATE	83
NOTE	74p.
AVAILABLE FROM	Information Reference Center (ERIC/IRC), The Ohio State Univ., 1200 Chambers Rd., 3rd Floor, Columbus, OH 43212 (subscription $8.00, $2.25 single copy).
PUB TYPE	Collected Works - Serials (022) -- Information Analyses - ERIC Information Analysis Products (071) -- Guides - Non-Classroom Use (055)
JOURNAL CIT	Investigations in Science Education; v9 n3 1983
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	*Academic Achievement; Attitudes; Cognitive Processes; Elementary School Science; Elementary Secondary Education; Environmental Education; Higher Education; *Inservice Teacher Education; *Participant Satisfaction; Performance Factors; *Preservice Teacher Education; Science Education; *Science Instruction; Science Teachers; Secondary School Science; *Teacher Behavior
IDENTIFIERS	*Science Education Research

ABSTRACT

Abstractor's analyses of 15 science education research studies are presented in three sections. The nine analyses in the first section focus on studies dealing with preservice and inservice teacher education. Areas investigated in these studies include: effectiveness of some inservice practices; teachers' perceptions of the importance of some teaching behaviors; questioning skills; teacher clarity; teacher vagueness; preservice training in environmental education; and local elementary school science practices. Four analyses of studies which examined factors affecting instruction are presented in the second section. Areas investigated include: students locus of control; thinking styles of students; self-concept and school achievement; the influence of mathematics ability on performance in physics; and student reading ability and the reading level of science materials. The final section contains a review and analysis of a single study. The study focused on attitudes and their influence on education in an informal setting. Authors responses to two analyses are also included. (JN)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED232869

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

SCIENCE EDUCATION INFORMATION REPORT

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
R.W. Howe

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

THE ERIC SCIENCE, MATHEMATICS AND ENVIRONMENTAL EDUCATION CLEARINGHOUSE
in cooperation with
Center for Science and Mathematics Education
The Ohio State University

SE042643

INVESTIGATIONS IN SCIENCE EDUCATION

Editor

Patricia E. Blosser
The Ohio State University

Associate Editor

Victor J. Mayer
The Ohio State University

Advisory Board

Robert L. Steiner (1983)
University of Puget Sound

Gerald Neufeld (1984)
Brandon University

Anton E. Lawson (1984)
Arizona State University

Lowell J. Bethel (1985)
University of Texas

Rodger Bybee (1984)
Carlton College

National Association for Research in Science Teaching

ERIC Clearinghouse for Science, Mathematics, and Environmental Education

INVESTIGATIONS IN
SCIENCE EDUCATION

Volume 9, Number 3, 1983

Published Quarterly by

The Center for Science and Mathematics Education
College of Education
The Ohio State University
1945 North High Street
Columbus, OH 43210

Subscription Price: $8.00 per year. Single Copy Price: $2.25
Add $1.00 for Canadian mailings and $3.00 for foreign mailings.

Volume 9, Number 2, 1983

NOTES FROM THE EDITOR iii

TEACHER EDUCATION 1

Glass, Lynn W. "Outcomes of an Energy Education Workshop for Secondary School Science Teachers." School Science and Mathematics, 81(6): 496-502, 1981.
Abstracted by THOMAS R. KOBALLA, Jr.. 3

Horak, Willis. "Behavioral Priorities of Science Teachers: An Analysis of Perceptions About Science Teaching." School Science and Mathematics, 80(6): 457-462, 1980.
Abstracted by HANS ANDERSEN 7

Riley, Joseph P., II. "The Effects of Preservice Teacher's Cognitive Questioning Level and Redirecting on Student Science Achievement." Journal of Research in Science Teaching, 18(4): 303-309, 1981.
Abstracted by EUGENE L. CHIAPPETTA. 11

Land, Michael L. "Actual and Perceived Teacher Clarity--Relations to Student Achievement in Science." Journal of Research in Science Teaching, 18(2): 139-143, 1981.
Abstracted by FRANCES LAWRENZ 15

Smith, Lyle R. and Grace H. Bramblett. "The Effects of Teacher Vagueness Terms on Student Performance in High School Biology." Journal of Research in Science Teaching, 18(4): 353-360, 1981.
Abstracted by THOMAS P. EVANS 18

Response by Smith to Evans's Critique 24

Towler, John O. "Observations on the Current Status of Environmental Education in Canada, Part I: A Survey of Canadian Pre-Service Training in Environmental Education." Journal of Environmental Education, 12(2): 11-16, 1980-81.
Abstracted by DAVID R. STRONCK. 26

Sheldon, Daniel S. and Dean Halverson. "Effects of a Televised Science Inservice Program on Elementary Teachers' Confidence Levels Concerning Science Concepts and Teaching Strategies." School Science and Mathematics, 81(5): 405-411, 1981.
Abstracted by CHARLES R. AULT, JR.. 30

Horn, J. G., and R. K. James. "Where are we in Elementary Science Education?" School Science and Mathematics, 81(3): 205-214, 1981.
Abstracted by WILLIS HORAK. 35

Voltmer, Rita K. and Robert K. James. "Laboratory Teaching Competencies for Science Teachers as Viewed by Science Educators." School Science and Mathematics, 82(3): 225-229, 1982.
Abstracted by DOROTHY GABEL. 38

FACTORS AFFECTING INSTRUCTION. 41

Schafer, John E. "Effect of Individualized Goal-Setting on College Biology Students' Locus of Control." Journal of Research in Science Teaching, 18(5): 397-402, 1981.
Abstracted by LAWRENCE L. GABEL. 43

Lynch, P. P. "Scientific Language and the High School Pupil: Preferential Thinking Styles." Journal of Research in Science Teaching, 18(2): 131-138, 1981.
Abstracted by RICHARD M. TOLMAN. 48

Campbell, Richard L. "Intellectual Development, Achievement, and Self-Concept of Elementary Minority School Children." School Science and Mathematics, 81(3): 200-204, 1981.
Abstracted by STEVE TIPPS. 56

Hudson, H. T. and Ray M. Rottman. "Correlation Between Performance in Physics and Prior Mathematics." Journal of Research in Science Teaching, 18(4): 291-294, 1981.
Abstracted by GLEN S. AIKENHEAD. 58

Response by Hudson to Aikenhead's Critique 60

Goldsmith, R. H. "Readability and Filmstrip Selection in Science." School Science and Mathematics, 80(4): 335-338, 1980.
Abstracted by JOHN EDWARDS 62

ATTITUDES. 65

Kiely-Brocato, Kathleen. "An Assessment of Visitor Attitudes Toward Resource Use and Management." Journal of Environmental Education, 11(4): 29-36, 1980.
Abstracted by RONALD D. SIMPSON. 67

NOTES FROM THE EDITOR:

This issue of Investigations in Science Education contains reviews of articles related to teacher education, both preservice and inservice; factors affecting instruction; and attitudes.

In the teacher education section, investigators looked at the effectiveness of some inservice practices (Glass, Sheldon and Halverson), teachers' perceptions of the importance of some teaching behaviors (Horak, Voltmer and James), questioning skills (Riley), teacher clarity (Land), teacher vagueness (Smith and Bramblett), preservice training in environmental education (Towler), and local elementary school science practices (Horn and James).

Among the factors affecting instruction that were investigated and reported are students' locus of control (Schafer), thinking styles of students (Lynch), self-concept and school achievement (Campbell), the influence of mathematics ability of performance in physics (Hudson and Rottman), and student reading ability and the reading level of science materials (Goldsmith).

The final section contains the review of a single article. The article was focused on attitudes and their influence on education in an informal setting (Kiely-Brocato).

Patricia E. Blosser
Editor

Victor J. Mayer
Associate Editor

TEACHER EDUCATION

Glass, Lynn W. "Outcomes of an Energy Education Workshop for Secondary School Science Teachers." School Science and Mathematics, 81(6):496-502, 1981.
Descriptors--Attitude Change; Energy; *Inservice Teacher Education; Knowledge Level; *Program Evaluation; Retention (Psychology); Science Education; Science Programs; *Science Teachers; Secondary Education; Secondary School Science; Secondary School Teachers; *Teacher Attitudes; *Teacher Workshops

Expanded abstract and analysis prepared especially for I.S.E. by Thomas R. Koballa, Jr., University of Texas at Austin.

Purpose

The purpose of this investigation was to determine the effectiveness of an inservice energy education workshop as indicated by high school science teachers' growth in knowledge and change in stated attitudes. The following questions were of primary concern in this study:

1. Does participation in the energy education workshop result in the cognitive growth of secondary school science teachers in the areas of: knowledge about energy resources, knowledge about the use of energy, and knowledge about energy conservation?
2. Does participation in the energy education workshop result in the stated attitude change of secondary school science teachers in the areas of: individual responsibility, collective citizen responsibility, corporate responsibility, governmental/legislative responsibility, and global responsibility?
3. Does cognitive growth and stated attitude change between pre- and posttest, if any, dissipate within one year following the energy education workshop?

Rationale

The research reported in this investigation is tied generally to the foreseen need to bring about curriculum change that provides an adequate information base and fosters judicious perspectives that will ensure survival in a world growing short of energy resources. Specifically, this study addresses the charge made by Boyer (1977) and the results reported by the National Assessment of Educational Progress (1977) by providing training for high school science teachers responsible for energy education curriculum development and instruction.

Research Design and Procedures

The study utilized a one-group pretest-posttest-delayed posttest design involving 25 inservice secondary school teachers selected from applicants to a National Science Foundation-sponsored energy education workshop. Pretest data were collected at the first workshop meeting and posttest data were collected at the last meeting. The delayed posttest data were collected by mail exactly one year following the completion of the 30 week workshop.

The Energy Inventory developed by the Biological Science Curriculum Study (BSCS) was used to measure eight relevant dependent variables. The inventory

consists of three knowledge subscales (energy resources, use of energy, energy conservation) and five attitude subscales (individual, collective citizen, corporate, governmental/legislative, global) that, when used collectively, total one hundred items. The questions addressed by this investigation focus on eight dependent variables that correspond to the eight subscales that constitute the Energy Inventory.

The workshop was organized into three separate components based on Fowler's (1976) hierarchical scheme for the development of energy education curriculum activities. Sources of energy was the focus of the first-third of the workshop; included were topics regarding fossil fuels, solar, and nuclear energy. The second-third of the workshop focused on energy demands; topics addressed were business, industrial, and agricultural uses of energy. The final-third of the workshop considered the various factors and forces affecting energy decision making.

Data collected from the use of the three knowledge and five attitude subscales were analyzed separately. In addition, data from related subscales were collapsed, permitting analysis of a multiple factor knowledge subscale and a multiple factor attitude subscale. A t statistic was used to test for significant differences between pre- and posttest means and between posttest and delayed posttest means.

Summary tables are reported for pretest and posttest comparisons and posttest and delayed posttest comparisons for the eight dependent variables.

Findings

At the conclusion of the 30 week workshop, significant mean differences were reported for the knowledge about the use of energy and the knowledge about energy conservation subscales. While neither the knowledge about energy resources nor a single attitude subscale reported significant differences between pre- and posttest means, both the total knowledge subscale and the total attitude subscale reported significant mean differences at the end of the workshop.

The energy knowledge and the stated energy attitudes had not changed significantly one year after the workshop from those reported at its conclusion.

Interpretations

As a result of this investigation it was concluded that energy education workshops can be designed that will result in relatively permanent knowledge acquisition and positive attitude shifts among teachers. Almost a 10 percent increase was reported for energy-related knowledge between the beginning and end of the workshop as measured by the total knowledge subscale. However, with mean scores on the pretest approaching a 75 percent correct response rate, changes much greater than those obtained on the total knowledge subscale could not have been expected. The increase in energy related knowledge was attributed to careful adherence to the workshop objectives that mirrored the knowledge subscales of the Energy Inventory.

Noting a pretest correct response rate of 55.3 percent for the total attitude subscale, it was anticipated that a relatively large attitude shift would

result if teaching occurred in a mannner that reflected the attitudinal objectives of the workshop. Only on the attitude toward individual response subscale did the mean attitude shift not correspond in relative magnitude to the total knowledge change reported. Possible explanations for the lack of changes on this subscale include: (1) the subscale items may not adequately reflect the intended attitudinal outcomes; (2) teaching may have been executed in a manner that did not correspond to the workshop objectives; and (3) requests for self commitment decreased the likelihood of changing personal beliefs and attitudes. Much remains to be done to circumvent the problems associated with personal commitment regarding energy related issues.

Participants' responses at the conclusion of the workshop to the question "What changes have been brought about in your classroom as a result of your participation in this project?" suggest that involvement in the workshop has had a major impact in their classrooms. The changes shared by the participants centered around the increased level of confidence in presenting current energy content to the students.

ABSTRACTOR'S ANALYSIS

There has recently been a lot of interest in studying the cognitive and affective advances of both teachers and students resulting from teacher participation in energy education symposia, workshops, and awareness programs. The investigator to this study addresses this timely concern.

This paper contains a representative review of the literature and develops a solid rationale for the investigation. While not of particular value to those attempting to build a model for knowledge acquisition or attitude change within the energy domain, it is worthwhile reading for persons charged with the responsibility for delivering energy education inservice programs.

In terms of the research design, there are some weaknesses that must be considered. The one-group pretest-posttest-delayed posttest design employed in this investigation failed to control for history, maturation, testing, or statistical regression and thus causes one to question the legitimacy of the suggested conclusions. The design does, however, provide some information about selection because the pretest describes the initial state of the selected participants on the eight dependent variables, but it falls far short of handling the other sources of internal validity. The shortcomings inherent in this design could have been thwarted by creating a "quasi-control group," accomplished by pairing comparable teachers with teachers participating in the workshop. Such an exercise in design gymnastics can be achieved by requesting that participants select a teaching colleague in the same school and grade to serve as a member of the control group (Bethel, et al., 1982).

Regarding the sample, the investigator speculates that the teachers selected to participate in the workshop were not significantly different from the population of science teachers who would normally attend similar functions. While this select sample may be representative of those who normally attend similar workshops, are they representative of the population of secondary science teachers responsible for energy education to which the results are generalized? A group of 25 hand-picked participants hardly represents an appropriate sample from which to draw generalizable conclusions. Furthermore, the sample of teachers selected as workshop participants represents an extremely

skewed sample as noted by their response rate on the Energy Inventory. With almost a 75 percent correct response rate on the pretest of the total knowledge subscales, it may be assumed that much of the information presented during the workshop was "old hat" to the participants. With participants needing little additional content, perhaps the focus should have been shifted more toward innovative implementation procedures. Concluding from the teachers' responses to the question regarding classroom impact, ideas for implementing energy information was what was truly desired.

The use of the Energy Inventory as the sole program evaluation instrument failed to account for all possible outcomes resulting from participation in the workshop. The instrument was reliable to the extent that the cognitive objectives of the workshop matched the items on the knowledge subscales of the inventory. However, as alluded to by the investigator, the affective objectives of the workshop which matched the attitude subscales were only marginally addressed. The lack of attention given to the attitudinal objectives clearly accounts for the less than expected mean attitude shift revealed by the data. Positive attitude shifts cannot be expected to occur solely as a result of the acquisition of additional knowledge.

Reporting the same instrument reliability information as the BSCS Instructor's Manual (1977), apparently the investigator made no attempt to establish the reliability of the Energy Inventory for the sample of workshop participants. Since sound results are inseparable from the soundness of the data producing instrument, failure to establish reliability for the instrument with the sample further causes the reported results to be questioned.

If replication of the investigation is desirable, concerted effort should be focused toward the attitudinal objectives if positive attitude shifts among participants is considered a worthy outcome. Also, components of the workshop that could have hampered positive attitude shifts need to be sought out and modified. For example, it is possible that some aspects of the workshop while adding to the participants' energy knowledge base could have been responsible for negative attitude shifts.

REFERENCES

Bethel, L. J., J. D. Ellis, and J. P. Barufaldi. "The Effect of an NSF Institute on Inservice Teachers' Views of Science and Attitudes Toward Environmental Science Education." Science Education, 66(4):643-652, 1982.

Biological Science Curriculum Study. Energy and Society: Investigations in Decision Making, Instructor's Manual. Northbrook, Illinois: Hubbard Scientific Company, 1977.

Boyer, E. L. "Special Feature on Energy and the Schools." Today's Education, 66(3):55-58, 1977.

Fowler, J. M. "Energy, Education and the 'Wolf' Criers." The Science Teacher, 43(3):25-32, 1976.

National Assessment of Educational Progress, Energy Knowledge and Attitudes. Denver, Colorado, 1977.

Horak, Willis. "Behavioral Priorities of Science Teachers: An Analysis of Perceptions About Science Teaching." School Science and Mathematics, 80(6):457-462, October 1980.

Descriptors--*Behavioral Science Research; Beliefs; Classroom Environment; *School Surveys; Science Education; *Science Teachers; *Statistical Analysis; *Surveys; *Teacher Attitudes; Teacher Characteristics

Expanded abstract and analysis prepared especially for ISE by Hans Andersen, Indiana University.

Purpose

The investigator's purpose in conducting this study was to assess science teachers' perception of the relative importance of eight specific categories of teaching behavior. The investigator also identified specific items with which teachers exhibited the most disagreement.

Rationale

Studies of teachers and teaching behaviors have been conducted by researchers with many different backgrounds. (Tyler, 1964; Mortenson, 1974; Bybee and Chaloupka, 1971; and Bybee, 1973, 1977, 1978). Many of the researchers have attempted to determine the characteristics of an ideal teacher. This investigation similarly sought to identify behaviors that science teachers perceived to be those of ideal science teachers. Knowledge of ideal teaching behaviors, it was assumed, would allow teachers to plan using more of the "ideal" behaviors and allow teacher education to prepare instructional materials and experiences that would facilitate the development of these behaviors among teachers.

Research Design and Procedure

A one-shot case study design was used. Eighty-six percent of the junior-senior science teachers from six school districts were invited to participate in the study and they comprised the study population. Demographic data were collected for each of the sixty-seven (67) participating teachers. Each participant then completed an 80 item Science Classroom Behavior Q-Sort (SCBQ) which was designed to measure the relative importance teachers placed on various classroom behaviors. The writer indicated the source of reliability and validity information (Horak, 1976). Each of the SCBQ categories contains ten items.

The major categories include: Warmth, Indirectness, Classroom Social System, Discipline and Behavior Modification, Management, Order and Control, Knowledge, Communication, and Flexibility and Variety.

The subjects sorted the 80 behavioral statements according to their beliefs about the importance of the classroom behaviors to science teaching. The items, so sorted, fell into an eleven point continuum from least to most important Q-factor analysis was used to analyze this data. Kendall's coefficient of concordance was used to determine the extent of agreement among the three types of teachers identified through factor analysis. Rankings of individual items were compared using Kendall's Coefficient to determine specific areas of agreement or disagreement. Disagreement was studied by comparing the overall standardized means of each of the 80 items.

Findings

1. As a result of using Q-factor analysis three distinct teacher types were identified.
2. Background variables were not important in the grouping of the teachers. The analysis used was not described.
3. Each teacher type held similar beliefs regarding their ranking of the relative importance of the categories. (Kendall's Coefficient of Concordance 0.70)
4. The rank ordering revealed the category of Communication had the highest teacher ranking and the categories of Discipline and Management ranked the lowest.
5. Computation of Kendall's coefficient of concordance did not indicate a significant association among the rankings of items in each of the categories which supports the claim that the categories are unique.
6. That different teacher types have divergent views about the nature of good science teaching was supported. A difference of more than one standard deviation was accepted as evidence of divergent opinion.

Interpretation

All three types of teachers in the study strongly expressed the opinion that being an effective communicator was the characteristic that was most important and that discipline and management was the least important determinant of effective teaching. Also evident was the fact, that while the three types of teachers agreed about the relative importance of these variables, they strongly differed in their response to individual items. The fact that three stable types were identified could permit schools to match teacher type to school type.

ABSTRACTOR'S ANALYSIS

The search for definitions, or a definition for a "good" science teacher has gone on for many years. In this article, the writer summarizes data collected from 67 teachers who were asked to rank the importance of eight categories of teacher behaviors.

Ranked most important to least important, these variables were:
(1) communication, (2) flexibility, (3) indirectness, (4) warmth, (5) knowledge, (6) social system, (7) discipline and (8) management.

An examination of these variables allows one to conclude that these science teachers rated as most important many of the same variables identified in the research synthesis prepared by Rosenshire and Furst (1971), which was subsequently supported by considerable research. The significance of this is that there is an agreement between what teachers believe is important to teaching, and data from studies that relate teacher behaviors to student achievement. Because interpretation of data from both types of studies tends to be parallel, both or either data could be used when considering changes in science teacher education programs.

In recent years I have spent considerable time observing practicing teachers who seem to rank knowledge first, and in reading research studies that indicate that teachers are again lecturing extensively and are again teaching science

as a rhetoric of conclusions. In this study teachers ranked knowledge fifth in importance. This is considerably different from my impression of what is happening in the classroom! Is this a unique group of teachers, or do teachers perhaps say one thing and practice another?

Perhaps the most significant finding was the fact that three different types of teachers were identified, defined, and described. Furthermore, the stability of each type was supported by applying Kendall's coefficient of concordance and obtaining a 0.70 value. Armed with data like these, schools should be able to assemble teaching teams with defined characteristics and begin studying the influence of such activities in student achievement.

The statistics used were appropriate and straight forward descriptive statistics. More statistical detail would be appreciated by many. However, this writer believes that the statistics are appropriate for the readers of the journal in which this article appeared. For example, when the investigator was examining response differences, he used one standard deviation as the indicator of significant difference when he could have used the Friedman two way analysis of variance by rank. (Siegal, 1956)

Searching for instructional variables that positively influence student achievement will continue to be useful research. The importance of determining how teachers rank teaching behaviors permits us to examine "how" they believe teachers should teach. If the "how to teach" can be related positively to student achievement and teachers use those behaviors they rank most important in their teaching, then we can use this type of easy-to-conduct research to guide our thinking about teacher training. However, more research of these potential relationships is necessary.

REFERENCES

Bybee, R. W. and D. W. Chaloupka. "Students' Perceptions of the Teacher They Like Best." Colorado Journal of Educational Research, 10(4):31-35, 1971.

Bybee, R. W. "The Teachers I Like Best: Perceptions of Advantaged, Average, and Disadvantaged Science Students." School Science and Mathematics, 73(5):384-390, 1973.

Bybee, R. W. "The Ideal Elementary Science Teacher: Perceptions of Children, Pre-service, and In-service Elementary Science Teachers." School Science and Mathematics, 75(3):229-235, 1975.

Bybee, R. W. "Science Educators' Perceptions of the Ideal Science Teacher." School Science and Mathematics, 78(1):13-22, 1978.

Flanders, N. A. "Teacher Influence in the Classroom." In E. J. Amidon and J. B. Hough (eds.), Interaction Analysis: Theory, Research and Application. Reading, MA: Addison Wesley, 1967.

Henney, M. and W. P. Mortenson. "What Makes a Good Elementary School Teacher?" Journal of Teacher Education, 25:312-316, 1974.

Horak, W. J. "An Analysis of Science Teachers' Beliefs About Teacher Classroom Behaviors." Paper presented at the annual meeting of the National Association for Research in Science Teaching. San Francisco, California, 1976, ERIC ED 124 424.

Rosenshine, B. and N. Furst. "Research on Teacher Performance Criteria." in Smith, B. O. Research in Teacher Education, A Symposium. Prentice Hall, 1971.

Siegal, S. Nonparametric Statistics for the Behavioral Sciences, McGraw Hill, 1956.

Tyler, L. "The Concept of an Ideal Teacher-Student Relationship." The Journal of Educational Research, 58(3):112-117, 1964.

Riley, Joseph P., II. "The Effects of Preservice Teacher's Cognitive Questioning Level and Redirecting on Student Science Achievement." Journal of Research in Science Teaching, 18(4):303-309, 1981.

Descriptors--*Academic Achievement; Elementary Education; *Elementary School Science; Elementary School Students; Higher Education; Preservice Teacher Education; *Questioning Techniques; Science Education; *Science Instruction; *Science Teachers

Expanded abstract and analysis prepared especially for I.S.E. by Eugene L. Chiappetta, University of Houston.

Purpose

The purpose of this study was to investigate the effects of asking lower and higher order questions and the redirection of questions on science achievement of primary and intermediate level elementary school children.

Rationale

The investigator pursued the line of research that attempts to determine how to improve retention and understanding of subject matter knowledge. He addressed two theoretical positions in this area of research: cognitive complexity and repetition.

One assumption is that if the instruction requires the learner to use higher level reasoning, he/she will remember more information because the interaction between the learner and subject matter is increased. For example, when the learner is asked to analyze and apply information, not only does he/she have to recall and demonstrate knowledge, but the students has to go through mental manipulations with the knowledge, which reinforces and further integrates it into the students' cognitive structure.

Redirecting questions in the classroom appears to be another reinforcing technique. When the same question is asked of more than one student, the other students are encouraged to attend and respond (either covertly or overtly) to the question. This keeps the students on their toes, and attending to the instruction. The redirection of a question also gives the students an opportunity to reflect on a given question, which may not occur when the teacher is delivering one question after the other in rapid fire order. Redirection improves wait time, which improves the quality and quantity of students' responses (Rowe, 1974).

Research Design and Procedure

Sample. A sample of 160 elementary school children was selected for this study. Five pupils were randomly selected from 16 intermediate (grades 4-5) classrooms and 16 primary (grades 1-3) classrooms.

Treatment. Each subject received one of three levels of questioning and one of two questioning strategies (redirection and no redirection). The treatments are as follows:

1. 100% High Cognitive Questions: Subjects in this group were taught from a prepared lesson script containing questions judged to be above the knowledge level (Bloom, 1956). The lesson was based on the time and location mini-lesson taken from the Teaching Improvement Kit (Popham, 1972). These kits contain mini-lessons with an explicit, measurable objective suitable for a 15 minute lesson.

2. 50% High Cognitive and 50% Knowledge Level Questions: Subjects in this group were taught from a prepared script containing 50% questions judged to be above the knowledge level and 50% at the knowledge level. The content for this lesson was also fom the same Teaching Improvement Kit mini-lesson used in treatment one.

3. 0% High Cognitive Level Questions: Subjects in this group were taught from a prepared script containing questions judged to be at the knowledge level and based on the same mini-lesson as those used in the other two treatments.

The subjects in each of the three treatment groups also experienced one of two questioning strategies.

A. Redirected Questioning: Subjects in this group had the same question asked by the teacher to two different students.

B. Directed Questioning: Subjects in this group had the teacher asking a question once and accepting the response before asking the next question.

Achievement Test. The achievement test consisted of 29 questions. Ten of the questions were at the knowledge level, fifteen at the comprehension level, and four at the analysis level. A panel of experts were used to judge the level of the questions, part of which were constructed by the author and part of which came from the Popham Teaching Improvement Kit. The total test reliability, using the split-half procedure was 0.91 and 0.94, 0.61 and 0.75 for the knowledge, comprehension, and analysis subtests respectively.

Research design. A 3 x 2 x 2 factorial design (question level x question strategy x grade level) with all factors fixed was used to analyze the data. MANOVA was used to test the significance of differences among treatment groups on the criterion measures and for interaction among the groups. The Pillai-Bartlett trace was used as the significance test.

Results

The author reported that there were no significant differences among the mean achievement test scores due to the cognitive level of questions used by the teacher. There was a significant ($p \leq .05$) difference among the mean achievement test scores at the knowledge level due to the effects of redirected questions asked by the teacher. There was a significant ($p \leq .05$) interaction that occurred between questioning level and questioning strategy on the comprehension substest and on the total test. Also, the intermediate grade students significantly ($p \leq .001$) out-performed the primary grade students.

Interpretations

The author reported that redirected questioning resulted in greater knowledge level achievement than nonredirected questioning. Asking the same question of two students, as opposed to one student, may result in an increase of student on-task involvement and reinforcement of learning.

ABSTRACTOR'S ANALYSIS

Improving science achievement is a worthwhile goal for science education researchers. This study adds more evidence to the research literature that achievement can be improved by questioning strategies that increase students' interaction with instruction. The abstractor believes that this line of research can provide even more evidence that higher order questioning as well as redirecting of questions can improve achivement. This can be accomplished by organizing the content of the instruction, the instructional strategies, and the text questions into identical categories of cognitive complexity-using Bloom's Taxonomy of Cognitive Objectives.

In order to achieve this uniformity, the lesson content, teacher questions, and test terms should be at the knowledge, comprehension, and application levels. If the education of the pupils is purposely targeted for these levels the outcomes may be pronounced. In the present study, this was not the case. The levels of questions asked by the teachers were probably mixed in two of the treatment (100% high cognitive questions and 50% high cognitive questions) groups. For example, in the treatment group that received 100% high cognitive questions, there were probably questions asked at the comprehension, application, analysis, and maybe synthesis and evaluation levels. Asking a mix of questions at these levels may not solidify elementary school pupils' thinking on a given topic, these pupils appear to need extensive cognitive experiences at a given level or with a given skill to master and reinforce their learning.

The transfer of training paradigm is one model used to help improve instruction and learning. If we are interested in analysis level learning for example, then a great deal of instruction at this level must be given. In the testing component, assessment must be at the analysis level; if the content is identical or similar in the instruction and testing, the test will be assessing knowledge or memory not analysis.

Matching content, instruction, and assessment presents another problem for the science education researcher besides balancing these factors in an experiment. Does the subject matter content lend itself to teaching and testing at the higher leveles of the taxonomy? Each science concept has a different structure, some are logically more suited to, say, testing at the application level, while other concepts can best be tested at the comprehension level. For example, electricity, atoms and molecules, flowers, and human circulation differ in their abstractness and complexity. At this point in the science education research history, it may make more sense to identify at which levels of cognitive complexity a given science concept should be taught and tested for a given age group of pupils. Some concepts may reach a learning plateau for a given age group at the comprehension level, while other concepts may be understood by the same age group at the application level. Unfortunately, the author of this study provided little information regarding the contents of the

time and location mini-lessons taken from the Teaching Improvement Kit (Popham, 1972), which were used in the instruction.

It might be productive for the researcher in this area to determine the amount of redirected question asking that is necessary to improve achievement for a given concept at a given taxonomic level for a given age group and developmental level of pupils. When teaching for the knowledge level learning maybe the amount of redirection will be strikingly different than when teaching for the application level. Teachers in the present study redirected a given question only once. And, of course, many other variables related to asking questions can be considered, such as the number of questions asked per session at each of the taxonomic levels, the type of wait time used, and the wait times recorded.

In conclusion, this study adds more support to the research literature that stresses the use of teacher redirected questions in the classroom Further research should identify science concepts for which instructional questions can be developed around at the knowledge, comprehension, and application levels, and for which test items can also be developed at the same levels as those used in the instruction. The research should also determine how much redirecting of questions is necessary at the various taxonomic levels to improve learning for a given age group of pupils and science concept under study.

REFERENCES

Bloom, Benjamin S. Taxonomy of Educational Objectives: Cognitive Domain. New York: David McKay Co., 1956.

Popham, W.J., E. Baker, J. Millman, and J.D. McNeil. Teaching Improvement Kit. Los Angeles: Instructional Appraisal Services, 1972.

Rowe, Mary Budd. "Relation of Wait-Time and Rewards to the Development of Language Logic, and Fate Control, Part II: Rewards." Journal of Research in Science Teaching, 11(4):291-308. 1974.

Land, Michael L. "Actual and Perceived Teacher Clarity--Relations to Student Achievement in Science." Journal of Research in Science Teaching, 1981, 18(2):139-143.

Descriptors: *Academic Achievement; *Classroom Communication; College Science; Higher Education; Science Education; *Science Instruction; *Student Attitudes; *Teacher Behavior; *Teaching Styles.

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Frances Lawrenz, Arizona State University.

Purpose

The stated purpose of this study was to determine the effect of three teacher clarity variables on student achievement in science and on teacher clarity as perceived by students. Four research questions were stated:

1. What are the joint effects of actual and perceived teacher clarity on student achievement?
2. What are the joint effects of level of achievement and actual teacher clarity on perceived teacher clarity?
3. What response items are most effective in distinguishing between clear and unclear lessons?
4. Is there any difference between the effects of high- and low-clarity lessons on achievement at the knowledge, comprehension, and application levels?

Rationale

This study was presented as providing experimental evidence for the effect of specific teacher clarity variables on student achievement. Other research in the field was cited but difficulties with it, i.e., its descriptive (correlational) nature and its use of non-operationally defined teacher behaviors, were pointed out. Two of the three clarity variables, mazes and use of "uh", were selected because previous research showed they were negatively related to student achievement in mathematics, but no research support for the inclusion of "OK" was given initially. No discussion of research on perceived versus actual teacher clarity was included.

Research Design and Procedure

The research involved randomly assigning 41 subjects to one of two treatment conditions, high teacher clarity (regular 18 minute lesson) and low teacher clarity (same 18 minute lesson interspersed with extra words (mazes), "uh's", and "OK's" at the rate of about 5 per minute for each). The lessons were videotaped and the same teacher presented each. The subjects were freshman and sophomore students in a general biology class. Immediately after viewing the lesson the subjects took a 24 item criterion-referenced content test (K.R.=.87) and completed a 15-item perceived teacher clarity instrument (take-retake r=.89). This is a randomized post test only design.

Findings

Several different analyses were competed to answer the proposed questions. The first analysis was a 2x2 ANOVA with actual and perceived clarity as the factors and student achievement as the dependent variable. The two actual

clarity levels were the two lessons, and the two perceived clarity levels were formed by grouping students scoring above or below the mean perceived clarity score. There were no significant effects for actual or perceived clarity and no significant interaction. The second analysis was a 2x2 ANOVA with achievement and actual clarity as the factors and perceived clarity as the dependent variable. The two levels of actual clarity were the two lessons, and the two achievement levels were formed by grouping students scoring above or below the mean achievement scene. The clarity main effect was significant but the achievement and interaction effects were not. Students rated the high clarity lesson higher than they rated the low clarity lesson. A third analysis was a t-test on whether students who perceived a clear lesson as clear achieved more than students who perceived an unclear lesson as unclear. This analysis produced significant results.

The third and fourth proposed research questions were answered without statistical analysis. The three items showing the highest mean difference between clear and unclear lessons in student ratings on the 15 item perceived clarity instrument were listed and are: well organized/not well organized, confident/not confident and well prepared/not well prepared. The fourth question was answered by listing the mean scores for the knowledge, comprehension and application portions of the 24 item achievement test for the high and low actual clarity groups. The differences were small .15 to .65, but favored the clear lesson.

Interpretations

The results of the study show that neither actual nor perceived teacher clarity (as defined by the three specific variables) has significant effect on student achievement.

The author suggests that the effect of teacher mazes may be different for different content areas because of the contradictory results of this study in science and previous research in mathematics. Alternatively, the author suggests that the use of "OK" may be positively related to student achievement in science.

ABSTRACTOR'S ANALYSIS

This investigation does add to the experimental evidence available on the effect of teacher variables on student achievement. As the author points out, little causal research is done in this area and it is necessary if definitive statements on how to teach are to be made. The research was carefully done so the results should be valid and the author presents his research in a clear, concise fashion.

The major shortcoming of this study is its scope. This is not necessarily a fault of the researcher but is related to this type of research in general. In order to have replicable studies everything must be carefully defined and controlled. Unfortunately this type of environment almost never occurs in actual teaching situations. Few teachers would say "uh", "OK" and extra words 5 times per minute throughout an entire class period, much less throughout a semester. Therefore, real students do not recieve treatments anywhere near that used in the experimental situation. Also, as the author points out, there are many aspects of teacher clarity that were not covered by this investigation.

Only relatively uncomplicated behaviors were included here, and these probably are not the ones that would figure most prominently in student achievement.

Even with the limited scope necessitated by experimental studies, this investigation could have been designed to answer more specific questions. For example, the use of three clarity behaviors rather than one at a time confounds the results--as the author discusses. Perhaps a series of smaller studies, each designed to answer one question and pave the way to the next, would have been more appropriate. (See Siegler (1976) for an excellent sample of this type of sequential research.) The first study could have followed the procedures used in Land and Smith (1979), where they found teacher mazes to significantly affect mathematics achievement, in a science setting to see if there were any subject effect. Then the use of "uh" and "OK" could be investigated independently to determine if both fuctioned as inhibitors of learning, and finally all three could be used in conjunction to determine the effect of the cluster. Additional experiments to test the congruence between actual and perceived clarity could also have been done. This type of research is much more time consuming, but it does somewhat make up for the lack of scope problem.

REFERENCES

Land, M. L., and Smith, L. R. "The Effect of Low Inference Teacher Clarity Inhibitors on Student Achievement." Journal of Teacher Education, 30:55-57, 1979.

Sigler, R. S. "Three Aspects of Cognitive Development." Cognitive Psychology, 9:481-520. 1976.

Smith, Lyle R. and Grace H. Bramblett. "The Effect of Teacher Vagueness Terms on Student Performance in High School Biology." Journal of Research in Science Teaching, 18(4):353-360, 1981.

Descriptors--*Academic Achievement; Biology; Botany; Communication Skills; Science Education; Science Instruction; Secondary Education; *Secondary School Science; *Student Attitudes; *Teacher Behavior; *Teacher Effectiveness; Teaching Styles.

Expanded abstract and analysis prepared especially for I.S.E. by Thomas P. Evans , Oregon State University, Corvallis.

Purpose

The purpose of the investigation was to determine the influence of teacher vagueness terms and the order in which the criterion instruments were administered on student achievement in high school biology and student perception of lesson effectiveness.

Rationale

Teacher clarity and its potential impact upon teaching and learning was the contextual framework within which the investigation by Smith and Bramblett was conceived. The investigation was the logical extension of previous writing and research by Struck and White (1972), Hiller, Fisher and Kaess (1969), Smith (1977), Smith and Edmonds (1978), and Land and Smith (1979). These references reported that teacher vagueness was one aspect of teacher clarity and that teacher vagueness influences student achievement in social studies and mathematics. Such findings suggested to the researchers of the present investigation that teacher vagueness terms may represent a global variable and influence student achievement across subject areas including science.

A further aspect of the investigation was related to an investigation by Smith and Land (1980) in which student perception of lesson effectiveness in mathematics was found to be negatively influenced by the use of a high frequency of vagueness terms by the teacher. The question was raised whether the results might not have been influenced by the order in which the test and teacher effectiveness questionnaire were administered. Smith and Bramblett decided to investigate this question.

A major underlying assumption in the investigation was that the achievement test, lesson effectiveness questionnaire, and teacher vagueness terms category system were valid and reliable.

A review by Taveggia (1974) was used to justify the use of videotaped rather than face-to-face lecture. Student achievement was found not to differ significantly as a result of the two teaching methods.

Research Design and Procedure

Forty-eight students from ninth grade honors biology classes were randomly assigned to one of four groups defined by the possible conditions of high versus low teacher vagueness and the order of administering an achievement posttest and lesson effectiveness questionnaire. The overall design may be

summarized as:

$$\begin{array}{llll} R & X_1 & O_1 & O_2 \\ R & X_1 & O_2 & O_1 \\ R & X_2 & O_1 & O_2 \\ R & X_2 & O_2 & O_1 \end{array}$$

where X_1 represents a high frequency of teacher vagueness terms, X_2 represents a low frequency of teacher vagueness terms, O_1 represents the achievement posttest, and O_2 represents student perception of lesson effectiveness questionnaire.

Twelve students in each group observed on of two 13-minute lessons on the vascular system of angiosperms. The lessons had been scripted, taught by the same instructor, and videotaped. Only the instructor's voice appeared on the videotapes along with overhead transparencies of diagrams and pictures. The tapes differed only in the frequency of teacher vagueness terms. The high teacher vagueness lesson (X_1) contained 140 vague terms while the low teacher vagueness lesson (X_2) contained five vague terms. The number of vagueness terms had been determined by analyzing each videotape using nine categories of vagueness terms as defined by Hiller, Fisher and Kaess (1969). The categories included: (1) Ambiguous Designation, (2) Approximation, (3) "Bluffing" and Recovery, (4) Error Admission, (5) Indeterminate Quantification, (6) Multiplicity, (7) Negated Intensifiers, (8) Possibility, and (9) Probability. Although individual terms had a higher degree of vagueness, all were given equal weight.

Student achievement and student perception of lesson effectiveness were determined by administering an achievement test and a lesson effectiveness questionnaire following the videotaped lesson. The achievement test consisted of 25 multiple choice and matching questions based on items from the textbook testing manual. The questionnaire required the students to rate the teacher from 1 to 4 on each of the following nine items: (1) The teacher was confident; (2) I was confident of the materials being presented; (3) The teacher was serious about the lesson; (4) The teacher's explanations were clear to me; (5) The teacher stayed on the main subject very well; (6) The teacher really knew what he was talking about; (7) The teacher did not seem nervous; (8) The teacher was prepared; and (9) The teacher did not seem lazy.

The data were analyzed by means of a 2 (low versus high teacher vagueness) X 2 (posttest then questionnaire versus questionnaire then posttest) analysis of variance. The analysis was conducted once using mean lesson evaluation scores for the total questionnaire and a second time using the mean scores of each of the nine responses to the questionnaire as dependent variables.

Findings

Student mean achievement and mean lesson effectiveness scores for the total of all nine questionnaire items were not significantly influenced by the order of administering the posttest and lesson effectiveness questionnaire.

Student mean achievement and mean lesson effectiveness scores for the total of all nine questionnaire items were not significantly influenced by the order of administering the posttest and lesson effectiveness questionnaire.

Student mean achievement and mean lesson effectiveness scores were significantly higher, p 0.05 and p 0.001, respectively, in favor of the low teacher vagueness condition.

The interaction between vagueness and order of administering the posttest and lesson effectiveness questionnaire was not significant in the analysis involving student mean achievement or mean lesson effectiveness scores for the total of all nine questionnaire items.

The mean lesson effectiveness evaluation scores for each of the nine questionnaire items were not significantly influenced by the order of administering the posttest and perception questionnaire.

The vagueness main effect significantly influenced ($p < 0.05$) the mean scores on 5 of 9 lesson effectiveness questionnaire items in favor of the low vagueness condition. The significant items included: (1) Item 1--The teacher was confident; (2) Item 3--The teacher was serious about the lesson; (3) Item 5--The teacher stayed on the main subject very well; (4) Item 7--The teacher did not seem nervous; and (5) Item 8--The teacher was prepared.

There was a significant interaction ($p < 0.05$) between vagueness and order of administering the posttest and lesson effectiveness questionnaire for item 7. Students in the high teacher vagueness condition rated the teacher as being more nervous if they filled out the lesson effectiveness questionnaire before they took the posttest.

Interpretations

The investigators cautiously concluded that vagueness items represent global variables that influence student learning and perception of lesson effectiveness across subject areas including secondary school biology. They further concluded that the order of administering the posttest and perception questionnaire does not significantly affect student achievement, or, with one exception, perception of lesson effectiveness.

A recommendation was made for those persons involved in teacher training and evaluation to place more emphasis on manageable skills that can be observed, quantified, and objectively critiqued. An example in the area of teacher clarity would be to assist teachers in reducing the use of vagueness terms in the classroom.

ABSTRACTOR'S ANALYSIS

These investigators have provided an example of research and research reporting with a number of characteristics that other science education researchers would do well to emulate. The design was simple, and, for the most part, the variables were well-defined and controlled. The report provided explicit statements about how the investigation was conceived and how it related and contributed to the existing body of research and writing. The report was clear and succinct. Although the teaching procedure used in the investigation was not similar to most educational situations, its use was substantiated by previous research. Several shortcomings of the investigation were identified, revealing considerable insight into the investigation and related research. The investigators were cautious in drawing their conclusions, and they pointed

out the reasons for the caution in the report. Plausible explanations were provided when the results of the investigation differed from findings or conjectures reached in earlier investigations.

One of the major drawbacks in science education research is the lack of an adequate number of and variety of valid and reliable criterion instruments. To produce a valid and reliable instrument requires considerable expertise and a great deal of time, effort, and resources. Even then the results are often disappointing. Because of these conditions, researchers frequently assume, accept the assumptions of others, or make some logical argument, that an instrument is valid and reliable. In each of these cases, the most common practice is not to field test an instrument before its actual use in the investigation. In fact, the application of an instrument in the investigation becomes a field test, and as likely as not, instrument analyses data are not included in the report. In some instances the investigation is completed and conclusions are drawn without even the investigator knowing for sure about the validity and reliability of the instrument, because the findings do not reveal any obvious clues. In other instances, the findings clearly indicate that there is some question about the instrument. The investigation by Smith and Bramblett provides an example of the latter. The findings reveal a potential problem with the validity and reliability of at least one of their criterion instruments--the 25 item achievement test.

The investigators reported that the mean score on the achievement test was 14.33 out of 25 points and cautioned about the test's overall level of difficulty when they drew their conclusions. Apparently the test was not previously field tested, although this information was not discussed in the report. Also, they did not report any analyses made of the test based on its use in the investigation such as item analysis or index of reliability. Regardless, the low number of test items and high level of difficulty, coupled with the brevity of the lesson, make it very unlikely that the test was either valid or reliable. First of all, it is not impossible but very difficult for such a short test to have high reliability. Secondly, generating 25 questions that measure short-term retention over 13 minutes of instruction requires a level of detail that would seem to jeopardize test validity and reliability. Thirdly and more importantly, the high level of difficulty reveals that the test may not have been reliable. The individual test items could not be at the 50 percent level of difficulty, the level necessary for obtaining the maximum degree of reliability when defined as internal consistency. A high level of difficulty means that the split-half technique would yield a low reliability coefficient. Since there are strong indications that the test had low reliability, its validity is questionable. An unreliable test does not provide valid information.

To conclude that the test was not valid or reliable with only the information provided in the report is obviously risky. The purpose of the discussion about how this conclusion was reached is not meant merely to criticize the investigators. In fact, their caution regarding the test was laudable. Other researchers might simply have avoided the issue. The intent of the discussion is to point out to other researchers the necessity for field testing instruments prior to their use in the investigation, performing instrument analyses in the field test and actual investigation, and including a summary of the analyses in the report. Without information regarding the analyses, it is difficult, if not impossible, to ascertain the merits of the findings and conclusions.

The use of the vagueness factor and associated categories is an extremently interesting technique for identifying and categorizing selected aspects of teacher verbal classroom behavior. It has a great deal of potential as a research and teacher training device. The technique and its application raises a large number of questions, some of which require additional information from Smith and Bramblett or the originators of the technique. Others suggest areas in need for further investigation by researchers in science education.

The first questions that come to mind relate to the category system and method of encoding the vagueness terms. How valid and reliable is the technique for identifying and classifying teacher vagueness terms? What is the method of encoding the behaviors? Does the category system have adequate objectivity; i.e., could independent observers, after a period of training, obtain satisfactory agreement with each other on the identification and encoding of vagueness terms? None of these questions are dealt with directly in the research report. Vagueness is defined, and the categories are listed along with the identified vagueness terms and their frequency of use under each category. The reference to Hiller, Fisher, and Kaess (1969), from which the technique was derived, does not provide a great deal of additional insight into the questions. They simply state that the vagueness factor was a psychological construct based on the authors' intuitive judgments alone. Perhaps the categories were developed in the same manner, but this is not evident from the reference. If these questions regarding validity, reliability or inter-observer agreement, method of encoding, and objectivity cannot be answered by Smith and Bramblett or the originators of the technique, then they should receive first priority in any future research involving teacher vagueness terms.

In the investigation and the reference to the technique, vagueness terms were given equal weights in the analyses of the data. This is no doubt the easiest and most common method used by those who research verbal classroom behavior, but does it seem logical that each vagueness term would be equal in its influence? Is a term classified as vague always vague, or does it depend in part on the context in which the term is found? How many vagueness terms are required per unit of time to significantly alter student achievement and perception of lesson effectiveness? Do certain patterns of vagueness terms become synergistic in their effects? These questions and others relating to the differential effects of teacher vagueness terms need to be further investigated.

Additional questions worthy of further investigation follow:

1. What is the frequency of use of teacher vagueness terms in regular instructional situations in science classrooms?
2. Is the frequency of use of these terms dependent upon identifiable variables such as teacher characteristics, grade level, and content areas?
3. What are the effects of these terms in regular instructional situations in science?
4. Do the terms have a differential effect on students with varying backgrounds, abilities, and interests?
5. Can science teachers be taught to significantly alter the frequency of vagueness terms they use in the classroom?
6. What are the most effective training programs for altering teachers' use of vagueness terms in science classrooms?

REFERENCES

Hiller, J. H., G. A. Fisher, and W. A. Kaess. "A Computer Investigation of Verbal Characteristics of Effective Classroom Lecturing." American Educational Research Journal, 6:661-675, 1969.

Land, M. L., and L. R. Smith. "The Effect of Low Inference Teacher Clarity in Inhibitors on Student Achievement." Journal of Teacher Education, 30:55-57, 1979.

Smith, L. R. "Aspects of Teacher Discourse and Student Achievement in Mathematics." Journal for Research in Mathematics Education, 8:195-204, 1977.

Smith, L. R., and E. M. Edmonds. "Teacher Vagueness and Pupil Participation in Mathematics Learning." Journal for Research in Mathematics Education, 9:228-232, 1978.

Smith, L. R. and M. L. Land. "Student Perception of Teacher Clarity in Mathematics." Journal for Research in Mathematics Education, 11:137-146, 1980.

Strunk, W., and E. B. White. The Elements of Style. New York: The Macmillan Company. 1972.

Taveggia, T. C. "Resolving Research Controversy Through Empirical Cumulation: Toward Reliable Sociological Knowledge." Sociological Methods and Research, 2:395-407, 1974.

IN RESPONSE TO THE ANALYSIS OF

Smith, Lyle R. and Grace H. Bramblett. "The Effect of Teacher Vagueness Terms on Student Performance in High School Biology." by Thomas P. Evans. Investigations in Science Education 9(3): 18-23 , 1983.

Lyle R. Smith
Augusta College

The review of Evans of the article by Smith and Bramblet (1981) was thorough, thoughtful, and insightful. A brief response to Evans' analysis follows.

Evans correctly states that the validity and reliability of the posttest is questionable. As stated in the article, it was assumed that the test would have adequate validity and reliability because the test questions were based on items in the testing manual that accompanies the textbook (Otto & Towle, 1977). However, the test scores were lower than the researchers expected, thus indicating that validity and reliability of the test may have been low. This problem indeed illustrates the need for pilot testing criterion measures of student performance.

Evans questioned whether vagueness terms can be observed and quantified objectively. Smith (1977), Dunkin (1978), and Dunkin and Doenau (1980) reported that coders can be trained to adequate inter-coder reliability to identify vagueness terms. Smith and Land (1981) reviewed research on teacher vagueness terms. Researchers interested in this variable should find the review helpful.

Evans noted that the vagueness terms and the categories of vagueness terms were given equal weights in the data analyses. In unpublished research recently completed, Smith addressed this issue. He found that vagueness terms in the "bluffing" and recovery category had less negative effect on student achievement than did vagueness terms in the categories of ambiguous designation, approximation, indeterminate quantification, multiplicity, possibility, and probability. However, this is the first of numerous studies that should be conducted to resolve this problem.

Evans asked several more questions concerning the vagueness variable. The answers to some of these questions follow. Descriptive research (reviewed by Smith and Land, 1981) indicates that teachers use an average of three to five vagueness terms per minute of teacher talk. The threshold at which vagueness terms impede learning has not been determined and may depend upon such student characteristics as abilities and interests. Smith and Edmonds (1978) reported that college mathematics student achievement was not affected significantly by lessons in which 3.6 vagueness terms per minute of teacher talk were used, whereas an average of 7.2 vagueness terms per minute of teacher talk had a significant negative effect on achievement. Dunkin and Doenau (1980) reported similar results for sixth grade social studies students in that 3.3 vagueness terms per minute of teacher talk did not affect achievement significantly, whereas 4.8 vagueness terms per minute significantly affected achievement. Student perception of lesson effectiveness appears to be more sensitive to teacher use of vagueness terms than does student achievement.

Evans also asked if teachers can be trained to significantly alter the frequency of vagueness terms they use in the classroom. Smith (1982) reported that such training is feasible. Teachers can be trained to identify vagueness terms by listening to tape recorded lessons they present. This process significantly reduced the number of vagueness terms the teachers used. Further research is necessary to determine the most effective programs for training teachers to use appropriate verbal classroom behaviors.

REFERENCES

Dunkin, M. J. "Student Characteristics, Classroom Processes, and Student Processes, and Student Achievement." Journal of Educational Psychology, 70:998-1009, 1978.

Dunkin, M. J. & S. J. Doenau. "A Replication Study of Unique and Joint Contributions to Variance in Student Achievement." Journal of Educational Psychology, 72:394-403, 1980.

Otto, J. H. & A. Towle. Modern Biology. New York: Holt, Rinehart & Winston, 1977.

Smith, L. R. "Aspects of Teacher Discourse and Student Achievement in Mathematics." Journal for Research in Mathematics Education, 8:195-204, 1977.

Smith, L. R. "Training Teachers to Teach Clearly: Theory into Practice." Paper presented at the meeting of the American Educational Research Association, New York, April 1982.

Smith, L. R. & E. M. Edmonds. "Teacher Vagueness and Pupil Participation in Mathematics Learning." Journal for Research in Mathematics Education, 9: 228-232, 1978.

Smith, L. R. & M. L. Land. "Low-inference Verbal Behaviors Related to Teacher Clarity." Journal of Classroom Interaction, 17:37-42, 1981.

Towler, John O., "Observations on the Current Status of Environmental Education in Canada, Part I: A Survey of Canadian Pre-Service Training in Environmental Education." Journal of Environmental Education, 12(2):11-16, 1980-81.

Descriptors--*Elementary Secondary Education; *Environmental Education; *Higher Education; Outdoor Education; *Pre-Service Teacher Education; Research Reports; State Programs; Surveys

Expanded abstract and analysis prepared especially for I.S.E. by David R. Stronck, University of Victoria, Canada.

Purpose

The purpose of this study was to survey the quantity and quality of Canadian pre-service training in environmental education.

Rationale

The Canadian constitution makes education strictly a matter for each province to govern, oversee and fund. Among the provinces, one may assume that there is a great variety of programs in the new curriculum field of environmental education. Davis in 1976 reported that environmental education existed in every province and territory with a wide variety of approaches and materials. Nevertheless, communication about environmental education in Canada is difficult because of the lack of a Canadian journal for environmental education and the lack of a Canadian national consensus regarding curricula. This study attempts to improve communication through describing the practices of preparing environmental educators by the teacher training institutions of Canada.

Research Design and Procedure

The design of the study was the one-shot case study, without reference to other studies or control groups. This pre-experimental design in the study assumed that the environmental education movement would have an impact on introducing pre-service training in environmental education into many teacher training institutions in Canada.

The procedure was to send a thirteen-item questionnaire to each of the 48 teacher training institutions in Canada. This survey sought data from the entire population of Canadian teacher training institutions, not merely from a sample of those institutions. The letter accompanying the questionnaire requested that the dean or department head of the appropriate section of the institution give the instrument to the faculty member who was most qualified to respond.

Findings

The response rate to the questionnaire was 85 percent. Only 43 percent (18) of the institutions offered a course in the methodology of environmental education during the 1977-78 academic year; 1104 students were enrolled in these courses. The major emphases in these courses were mainly ecological, outdoor education, or biological in that rank order. Some of the courses concentrated on the methodology of science, social studies, or outdoor education and included little or nothing on environmental education.

Sixty percent (25) of the institutions offered a non-methods course dealing with substantive content relating to ecology or environmental concerns. In the 1977-78 year, 1210 students were enrolled in such courses.

There were 33 full time and 31 part time faculty members teaching environmental education. Only five of these instructors had an academic background in environmental education; most had degrees in education, biology or science, in that order. Eight percent (5) of these faculty members were involved in funded projects on environmental education research. Thirty percent (20) were designing curriculum materials in this field.

The respondents reported that environmental education was one component of the social studies area in the elementary grades and an element of the science program at the secondary level. Another subject area which frequently contained environmental education components was outdoor education.

The provinces do not give teacher certification in environmental education. Seventeen of the respondents favor such certification while 15 are opposed.

The major problems concerning the teaching of environmental education in these institutions were primarily lack of funding and of communication among environmental educators. Other major problems were texts for teachers, teaching materials and equipment, lack of research, and lack of Canadian content. Twenty percent (9) of the respondents were able to name an exemplary project in their region.

Interpretations

The investigator concludes that "it will be some time before there will be a national consensus on the importance of preparing teachers especially trained to teach environmental education." This new field lacks attention and publicity. Few professionals are involved in research in this field or in communicating with their colleagues. Few of the faculty members teaching environmental education have adequate academic preparation for this responsibility. The investigator also observes: "It is a strange anomaly that more institutions and students are involved in courses dealing with environmental content than with methodologies for how to teach this content."

ABSTRACTOR'S ANALYSIS

The investigator did an original study in an area that has been generally neglected. Two previous studies (Rioux 1973 and Davis 1976) had considered environmental education in Canada, but emphasized classroom practices at the pre-college level. This survey on pre-service training provides helpful data.

At the beginning of the questionnaire, the investigator provided a long and complete definition of environmental education that is both correct and rather universally accepted. Unfortunately the clarity of this definition does not characterize the items in the questionnaire. For example, the first question is "Does your institution offer a course in the methods of teaching environmental education?" Forty-three percent (18) of the respondents said "yes" and provided course descriptions or outlines. Many of these course descriptions or outlines revealed that the courses had little or almost nothing to do with environmental education. A definition of "a course in the methods

of teaching environmental education" seems needed in order to avoid ambiguities. Without such a defined standard, the reader knows only that less than 43 percent of the institutions provide a course in the methods of teaching environmental education.

The investigator wrote: "It is a strange anomaly that more institutions and students are involved in courses dealing with environmental content than with methodologies for how to teach this content." A reader would not find any anomaly here because the item in the questionnaire simply asks: "Does your institution offer a course involving content pertaining to ecology or environmental concerns?" The respondents were probably describing the fact that universities offer courses in environmental studies that include more students than those who are seeking certification as teachers.

The last item on the questionnaire could not be interpreted because of lack of definition. The item was "If you know of an exemplary project or programme in environmental education in your area, would you please list it together with the name and address of a person whom we may contact for additional information." The investigator observed: "Only about 20 percent of the respondents were able to do so. The reasons for this are unclear, and it is not possible to determine whether this is an aberation in the results, a true picture of a terrible lack of activity in the field, the outcome of unreasonably high standards applied to the term 'exemplary,' or a variety of other factors." The abstractor believes that British Columbia has a very large number of exemplary projects or programmes in this field. Perhaps many of the respondents were reluctant to provide names and addresses because few of the leaders in these programmes wished to be contacted for additional information.

The investigator found that lack of communication ranks at the top of the list of problems concerning the teaching of environmental education. He also noted that while there is no Canadian journal of environmental education, there is an historical reluctance to rely on non-Canadian periodicals and materials. The abstractor further observes that other barriers to communication are the great diversity of environmental problems extending across the vast geography of Canada and isolation of educational activities within each province because of the Canadian constitution. Probably the investigator encourages the use of The Journal of Environmental Education as an appropriate means of communication for Canadian educators because he published this article in that journal.

The investigator is probably very correct in concluding that a national consensus on the importance of preparing teachers especially trained to teach environmental education will be long delayed. There seems to be resistance against accepting even basic definitions. For example, in 1980 the Ministry of Environment in British Columbia produced the Environmental Education Handbook which ignores classic definitions and encourages confusion by the following judgment.

> "Environmental and Outdoor Education are two terms very commonly used in this booklet. We do not distinguish between them because in many ways they are synonomous."

At the University of Victoria, courses designed to prepare teachers for environmental education were described as "outdoor education." During recent years the economic recession has led to the elimination of many courses in

environmental education at pre-college level and some in the pre-service training institutions of Canada. If the survey were replicated in 1983, the new data might demonstrate a serious decline in activity in Canadian environmental education.

REFERENCES

Davis, J. G. Environmental Education in Canada - 1976. Toronto: Ontario Ministry of Education, Curriculum Branch, 1976.

Rioux, J.C. Environmental Education in Primary and Secondary Schools in Canada. Ottawa: Environment Canada, 1973.

Sheldon, Daniel S. and Dean Halverson, "Effects of a Televised Science Inservice Program on Elementary Teachers; Confidence Levels Concerning Science Concepts and Teaching Strategies". School Science and Mathematics, 81(5):405-411, 1981.

Descriptors--Educational Television; Elementary Education; *Elementary School Science; Evaluation; *Inservice Teacher Education; Science Course Improvement Projects; Science Education; *Science Instruction; *Science Teachers; Teacher Characteristics; Teaching Methods; *Telecourses; Television

Expanded abstract and analysis prepared especially for I.S.E. by Charles R. Ault, Jr., Indiana University

Purpose

The investigators sought to determine the success of a district inservice program consisting of five televised 30 minute films designed to instruct elementary teachers on science concepts and teaching strategies selected from the SCIS curriculum newly adopted by the district. The criterion of success was improvement on a measure of confidence level.

Rationale

Television has many attractive features from the point of view of administrators planning total district, inservice programs: easy access, elimination of travel time and expense, well-rehearsed presentations, and the potential for repeated usage. As a relatively new mode of inservice teacher training and in consideration of its particular advantages, televised inservice programs merit careful evaluation. The need for science educators to attend to the problems encountered by elementary teachers in regards to grasping the often difficult or unfamiliar content and adopting the non-conventional teaching practices embodied in elementary science curricula such as SCIS underscore the importance of determining effective means of inservice training. Jones (1973) and Piper have found television effective when used in extended (fifteen or more presentations) inservice projects. Jones (1973) claims that comparisons of extensive televised to face-to-face presentations fails, in most cases, to establish significant differences. The question arises as to the minimal amount of televised inservice training expected to bring about changes in teacher knowledge of science concepts and teaching strategies.

The criterion used by the investigators to determine change in teachers' understandings of SCIS concepts by grade level and of science teaching strategies was change in score on an instrument assessing confidence levels. Provided that teacher reported confidence levels do, in fact, represent degrees of facility with concepts and teaching strategies, this procedure was valid. This relationship was the central assumption of the study and derives from Crumb's (1977) work on assessing televised inservice-training for science teaching.

Research Design and Procedure

The sample consisted of one hundred thirty-eight teachers from the elementary schools of Dubuque, Iowa attending inservice-training sessions for all the

elementary teachers in the district. The district had adopted the revised Science Curriculum Improvement Study (SCIS, 1978) for the 1978-1979 academic year. Using a one group, a pretest-posttest design, a SCIS Concept Inventory (adapted from Crumb, 1977) and a Science Teaching Attitude Scale (STAS, Moore, 1973) were administered to each of the one hundred and thirty-eight participants before and after the televised sessions. In addition, participants evaluated the workshop upon its completion using the Workshop Evaluation developed by Halverson (1979).

The televised sessions utilized ten video tapes from the commercial series, Science in the Elementary Schools developed with NSF support, as well as locally produced materials. Three programs, reviewed by all the teachers, focused on science process skills and pedagogy (the "SCIS learning cycle" for example). The remainder of the programs centered on presentations of SCIS physical and life science concepts by grade level. Teachers were required to view two of these programs, bringing the total number of inservice sessions for an individual participant to five. Treatment varied, therefore, by grade level. Only the first program was aired on an inservice day; all others were scheduled at the end of regular school days. The inservice coordinator also distributed and discussed four papers with the participants prior to the televised sessions.

> "Piaget was a Quarterback" (Avdul, 1974),
> "Science, Silence, and Sanctions" (Rowe, 1975),
> "NSTA Position Statement on School Science Education for the 70s" and,
> Individualization through the Laboratory Approach" (Lavatelli and Thier, 1968).

The SCIS Concept Inventory assessed progress in understanding of concepts and teaching strategies in terms of confidence levels judged according to the following:

A. Fully knowledgeable enough to design and teach a series of lessons using the topic.

B. Know how to define and explain the meaning adequately, but could not use it in designing and teaching a series of lessons.

C. Have only a vague, superficial, and limited understanding.

D. Have poor or no understanding at all.

Using these criteria, the participants responded to fifteen SCIS teaching strategy concepts and ten science concepts specific for each grade level.

To infer significant changes at the .05 level in teachers' knowledge of science concepts and teaching strategies, t-tests were made comparing mean pre- and posttest concept inventory scores (= confidence levels). Scores were reported by grade level.

Findings

Setting the level of significance at .05, the authors found significant improvement in teachers' knowledge of science concepts and teaching strategies

for most grade levels as determined by the t-test comparing pre- and posttest means on the instrument designed to measure changes in confidence levels. The data indicated that teachers in grades 2, 4, 5, 6 significantly improved on science concept scores. (A disagreement between the data in the authors' Table I and their reporting of significance values in the text raises questions about the grade 2 and 3 results). Gains at all grade levels were similar, despite initially wide ranging pretest scores. Teachers at all grade levels significantly improved on understandings of SCIS teaching strategies according to the confidence level criterion. "Seventy-three percent of the participants indicated that future inservice programs should be conducted using television." By a three to one ratio, participants preferred the locally provided materials as compared to the commercially produced programs.

Interpretations

The authors claim that as few as five televised thirty-minute inservice presentations are effective in improving teachers' understandings of SCIS science concepts and teaching strategies. The authors are quick to point out that the criterion for this claim is gain in a measure of confidence.

The preference for local productions suggests the value of "individualism" (appropriateness for a particular set of teachers) of those segments.

ABSTRACTOR'S ANALYSIS

The authors are to be applauded for reporting on the implementation and evaluation of an inservice workshop in a manner which is both replicable and informative. They have attempted to infuse the skills of research into the assessment of the efficacy of televised inservice training. Their questions in regard to this goal are clearly stated.

The criticisms of the research come under four headings:

1. Validity of confidence levels as a criterion of understanding.
2. Generalizability of the findings given the design and treatment.
3. Discrepancy/confusion on treatment procedure and data reporting.
4. Substantive comments by participants on the effects of the televised segments.

Confidence levels as a criterion of understanding

The authors are careful to embed their work in related studies using confidence intervals as the criterion of understanding. Obviously, the entire value of the study hinges on the inference of gain in knowledge of science concepts and improvement (or change) in teaching strategy from confidence level score changes. The SCIS Concept Inventory (adapted from Crumb, 1977) apparently achieves respectable split half reliability (.66). However, in the context of testing a minimal approach to bringing about desired change in teacher knowledge and skill, the conclusion that such an instrument, in fact, measures more than confidence gain is suspect without employment of additional, cross-validating procedures.

The authors did report that Crumb's (1977) format "when used as a posttest with control and experimental groups" yielded a "control group . . . mean score of 70 as compared with a mean score of 120 for the experimental group". Since the nature of control and experimental groups are not described, the meaning of this difference is unclear.

In effect, the authors embed their most important assumption neither in a theoretical tradition reducing doubt nor in an adequate web of substantiating empirical work. Their two-part justification for testing knowledge in an unconventional sense is based on practical considerations but remains unconvincing. They argue for the value of an instrument "less threatening to teachers because of the absence of right and wrong answers".

The assertation that the confidence level expressed for a particular concept measures the supposedly important "facility with the concept" is vague. What does "facility" mean; how would one make a record of it? When asking important questions about methods of improving teachers' knowledge of science concepts, researchers ought not to retreat from substantively finding out what teachers know and how they actually use that knowledge. We need to know whether teachers are entitled to their confidence-or unjustified in their lack of it.

Design and treatment limitations to generalization

The limitations of a one-group pretest-posttest design obviously detract from the validity of this study. As a matter of practical and ethical necessity, all inservice participants were treated the same. A demonstration of the positive effects of the televised training was, of course, the desired and obtained outcome. Few grounds were offered for generalizing the findings to other settings and internal weaknesses in the design and treatments (particularly the use of confidence level as the criterion of understanding) further weaken the authors' claims. What role, for example, was played by the several articles which provided background to the training? Were effects related to the style and content of the televised segments or to participants' interaction during the time between sessions? How much improvement in confidence was due to the participants' expectations that they would understand SCIS better? The design permitted no control over these factors.

No demographic data were reported, though these were collected. No background differences in prior science and science methods training were mentioned, an area of interest to explore in regard to the range of scores.

Discrepancy/confusion

Unclear to this reader was the relationship among two crucial instruments to each other and to the pretests and posttests. The authors reported using their own SCIS Concept Inventory adapted from Crumb's (1977) format and Moore's (1973) STAS. Neither the data table titles nor the text discussion adequately described how scores were derived from these two instruments, particularly the teaching strategy scores.

The data table and text conflict as to the significance level of pre- and posttest means for second and third grade teachers on science concepts. The discrepancy is frustrating.

Comments by participants

Despite informing the reader as to the overwhelming positive evaluation given by teachers to the workshop, no substantive information was presented giving clues as to what aspects of the instruction were appreciated. How the content of locally produced segments was tailored to local, inservice needs might be of interest to other administrators contemplating a similar inservice program.

Closing remarks

Sheldon and Halverson's study hints at the promise of televised inservice training programs. They document significant changes following a minimal series of workshop events. A more penetrating look at the nature and depth of those changes is desirable. The utility of confidence levels as a measure of understanding merits further inquiry.

REFERENCES

Avdul, R. N. "Piaget was a Quarterback." Teacher, 81(8):10-12, April, 1974.

Crumb, Glenn H. "Using Mass Media for Inservice Teacher Education." A report presented to the Council of State Science Supervisors. Western Kentucky University, Bowling Green, Kentucky, 1977.

Halverson, Dean L. "The Effectiveness of a Televised Science Inservice Model for Improving Elementary Teacher Attitudes and Concept Knowledge." Doctoral Dissertation, University of Iowa, Iowa City, Iowa. May, 1979.

Jones, R. M. "An Application of System Analysis Techniques in the Development of Instruct. Tele. Units of Inservice Elementary Teachers." Norman, Oklahoma: Oklahoma State University, 1973 (ERIC Document Reproduction Service No. ED 083 832).

Lavatelli, C. B. and H. B. Thier. "Individualization through the laboratory approach." SCIS Newsletter, University of California, No. 12:4-5, Spring, 1968.

Moore, R. W. "The Development, Field test, and Validation of Scales to Access Teachers' Attitudes Toward Teaching Elementary School Science." Science Education, 57:271-278, 1973.

Piper, Martha Kime. "The Development and Evaluation of a Televised Science Inservice Program." Ph.D. Dissertation, University of Texas at Austin, 1974.

Rowe, M. B. "Science, Silence and Sanctions." Science and Children, 6(6):11-13, March, 1969.

Horn, J. G., and R. K. James "Where are we in Elementary Science Education?" School Science and Mathematics, 81(3):205-214, 1981.

Descriptors--Elementary Education; *Elementary School Science; *Questionnaires; Science Course Improvement Projects; Science Curriculum; Science Education; Science Instruction; *Teacher Attitudes; *Teacher Background; *Teacher Characteristics; *Teacher Qualifications; Teacher Response

Expanded abstract and analysis prepared especially for I.S.E. by Willis Horak, The University of Arizona

Purpose

The study reported was designed to answer questions relative to local practices in elementary science education. It was designed to parallel three national studies and to additionally reveal local or state idiosyncracies in Kansas. Overall, the study investigated 12 basic questions relevant to science teaching in elementary schools. Specifically, these questions related to teacher background information, to teacher preparation and practices, and to perceived teacher needs.

Rationale

The rationale for this study was the identified need to know exactly what kinds of elementary science instruction children in the State of Kansas were receiving. Additionally, there was the need to know if the local or state problems were indeed those identified and discussed by the three previous national studies. The areas of need and support could supposedly help local school districts and colleges/universities to develop programs that would enhance the teaching of science in the elementary schools. Through first assessing and analyzing factors that influence elementary science teachers, effective training programs could be developed which would capitalize on areas teachers perceived to be the most relevant. Thus, the rationale for the study was based mainly on the need for base line data which professionals in the State of Kansas could use to expand and update elementary science instruction.

Design and Procedures

Since the purpose of the study was only to gather and analyze base line data, the investigation was basically a one time case study. A questionnaire which reflected both local and national concerns in science education was developed. This questionnaire was then mailed to a random sample of teachers drawn from the records of the Kansas Department of Education. The questionnaire contained 12 basic areas of concern. The collected data were analyzed separately for each area of concern. Profiles were established which revealed in tabled form the assessed results for the areas. The means and standard deviations were computed for comparison purposes where applicable.

The return rate of the questionnaire was very poor. Overall, only 22 teachers responded during the time of the study. The authors felt justified in reporting the results from such a small sample size since the findings signified the trends which additional data may have more fully revealed. In addition, they state that the local results were consistent with the findings in the larger

national studies and that after an effort was made to gather and compare responses from non-responders with the responses from responders, no major differences could be identified.

Findings

The results of the study were summarized in twelve separate tables. The breakdown of the overall background characteristics of the respondents were similar to those typically found in the Midwest. Most of the elementary teachers were female and did not hold a degree beyond the BA level. Additionally, their undergraduate area of concentration was in some subject other than science or mathematics. They thus did not feel prepared or qualified to help other teachers in any of the areas of elementary science instruction. One other item of importance concerned the fact that only two of the respondents had any participation in an NSF sponsored institute, conference, or workshop. They consequently reported that they made limited use of the major "project" curricula.

The science teaching practices of the respondents varied greatly. Class discussion was the most frequently used technique for teaching science. Hands-on manipulatives or laboratory materials were not generally used daily but were most often used at least once a week. When comparing the time of instruction devoted weekly to teaching science, mathematics, social studies, and reading it was ascertained that science instructional time ranked last. Lastly, most teachers had available portable science kits or materials to use in their classrooms. However, a few (15 percent) reported that they had no facilities or materials available to them.

When responding to items relating to factors affecting science instruction and additional teacher needs, it was reported that time demands on teachers, attitudes toward science and inadequate teacher preparation were most important. Teachers most needed help using and maintaining manipulatives and equipment. They felt storage space, work space, and money to buy supplies were hindering them from providing good science instruction.

Interpretations

This study pointed out that local and state elementary science teachers in Kansas share many things in common with those assessed in the national surveys. Effective science training is needed for many teachers. Through assessing local needs, we may best stress the things teachers feel are important. Society and teachers must value scientific literacy if we are to succeed in offering our elementary children more and better science programs.

ABSTRACTOR'S ANALYSIS

In the introduction to this article the authors state that, although three major studies have been carried out on a national level, they are not sure what schools are doing in elementary science at a local level. Thus, in their study they tried to provide specific answers for the State of Kansas. I am sure many science educators read the national studies and had similar thoughts. Therefore, this study was worthwhile and useful in that it provided readers an example of how local perceptions and practices may be related to the published national ones.

When one compares the results obtained in this study with the national results, one sees many similarities. In fact, overall, there are few discrepancies in areas where both studies assessed the same state of science instruction. This allows one to more fully appreciate the generalizability of the national studies. It also allows persons to assess areas other than those already assessed by the national studies if they want to get baseline information about various aspects of elementary science education. This is a great advantage since one is often tempted to make the questionnaires too extensive.

The overall validity of the results of this study may be questioned due to the small sample size. As previously stated, the authors do make a case for analyzing the results even though there were only 22 respondents. Part of the explanation is, however, vague. It is stated that ". . . a comparison of responses between respondents and non-respondents revealed no significant differences . . ." What is a response of a non-respondent? Did the authors contact some non-respondents by telephone or by mail and get them to respond? Did they simply contact some other teachers who were not in the original sample and compare these answers with those of the original 22 respondents? This point needed a more thorough explanation. I also feel that the choice of descriptive words "no significant differences" in part of the sentence previously cited is somewhat misleading. In education when we talk about significant and non-significant differences, we are usually referring to some type of statistical test at a specified level of significance. This does not appear to be the case for this study.

The information gained in this study and presented in the tables is worth knowing and useful. It would, however, also be helpful if an attempt was made to answer questions about possible interrelationships in the tables. For example, if we consider Tables I and IV, we might consider whether teachers in the lower elementary grades use hands on manipulatives and/or laboratory materials more, less, or the same often as teachers in the upper elementary grades. Similarly, for Tables I and VI we might also try to determine whether the 45 percent of the teachers with the MA, MS, or MAT degrees teach science for a larger or smaller amount of time than those teachers whose highest degree is the BA or BS. The small sample size would adversely affect the types of questions that may be answered. In cases where there are minimal amounts of reported data, we would have to be extremely careful in making broad generalizations.

After comparing this study with studies of similar nature, one concludes that in the future we need to direct research studies to do more than gather baseline information. We need now to specifically determine ways to ameliorate the problems identified. These ways may be delineated by more comprehensively seeking relatinoships among the answers to the posed questions. If we do find that teachers use more inquiry methods or laboratory procedures when dealing with science topics they perceive to be relevant, then we may devise plans of study that utilize this knowledge. Similarly, if we find that teachers do not value scientific literacy, whereas they do apparently value mathematics literacy, then we must devise studies to determine why this is true.

Finally, science educators must continue to seek information relative to the status of science in our elementary schools. They must continue to analyze and report such data. This study thus provides them with additional information. It is now the task for future researchers to devise studies which determine how best to effect change in the scientific attitudes and practices of this generation of elementary teachers.

Voltmer, Rita K. and Robert K. James. "Laboratory Teaching Competencies for Science Teachers as Viewed by Science Educators." School Science and Mathematics 82(3):225-229, March, 1982.

Descriptors--Educational Research; Higher Education; *Laboratory Training; *Preservice Teacher Education; Science Curriculum; *Science Education; *Science Instruction; *Science Laboratories; *Teacher Education.

Expanded abstract and analysis prepared especially for I.S.E. by Dorothy Gabel, Indiana University.

Purpose

The investigators' purpose for conducting this study was to determine the appropriateness of certain laboratory teaching competencies for inclusion in secondary science education programs and the extent to which students had developed such competencies as viewed by science educators.

Rationale

Because many teachers do not possess laboratory teaching competencies to successfully use a laboratory approach in teaching science, laboratory competencies should be identified and taught to secondary preservice teachers. This study attempts to identify these competencies.

Research Design and Procedures

A list of 70 laboratory teaching competency items was distributed to 110 college and university science educators from across the United States in this survey study. The items had been refined from feedback attained on 85 items from a small sample of science educators. Seventy-five percent of the original science educators, who were carefully selected according to criteria of direct association with teaching secondary preservice teachers, completed the survey. They rated each of the 70 competencies on a nine point Likert-type scale ranging from Low to High according to the appropriateness and the level of student competence.

Findings

Results indicated that science educators found eleven of the 70 competencies inappropriate for inclusion in a science teacher education program. These included the operation and application to the science curriculum of the oscilloscope, spectrophotometer, Geiger-Muller counter-ratemeter, helium-neon laser, kymograph, microtome, sphygomomter, photographic film development, production of slides and prints, stoboscopic techniques for time and measurements, and introduction to and the uses of the planetarium.

These items were also rated low in terms of students' knowledge of their use in addition to 20 other items.

Interpretations

Science educators viewed approximately 85 percent of the 70 laboratory competencies as appropriate for beginning science teachers. Items perceived

as important to be included in the curriculum were also found to rate high in terms of student competence. The list of competencies might be a helpful guide in supplementing the work of practicing science teachers and preparing inservice teachers.

ABSTRACTOR'S ANALYSIS

This study addressed the very practical question of what competencies are appropriate for secondary science teachers as viewed by science educators. The question is important because these science educators determine which competencies will be incorporated into preservice science methods courses.

The major strength of this study is in the selection of science educators for participation in the survey. Great care was taken to set up a rigid qualification system that was applied in their selection. In addition, the 75 percent return rate of questionnaire was quite satisfactory.

There are several facets of this study that need explication and clarification to make the research report a valuable tool for science educators. The first of these is the validity of the list of competencies considered to be inappropriate for preservice teachers. Most of the competencies listed as inappropriate are related to the teaching of physics or the physical sciences. The question must be asked as to whether the science educators were rating competencies for teachers of general science or competencies for specific teaching areas such as biology, chemistry, and physics. Students generally major in a specific science area. It would seem more valuable to have obtained lists of competencies according to specific science disciplines. Another question that must be asked is the background of the science educators. A science educator with a biology background might rate competencies differently than an educator with a chemistry background. Were science educators with the various disciplines equally distributed? If not, the results may be biased.

Another question concerning these competencies is whether or not science educators are the appropriate persons to make judgments on competencies. It would appear that practicing teachers could make a valuable contribution to the research in this area.

The question of the reliability of the instrument is also of some concern. Using a Likert-like scale with a wide range from 1-9 involves making rather fine distinctions about the value of the competency. No reliability coefficients are given in the research report. It would be interesting to know if science educators would rate these items in an identical way a second time. A Q sort might have been the more appropriate procedure to use for the rating system.

One of the deficiencies of this research report lies not in the research itself but the failure of the report to list the competencies that were found to be important. It would be desirable to include these because it would be more valuable to the science educator to know what to teach than what to avoid teaching. Although these would be available from the authors, it is always an inconvenience to write for more explicit information that would not have taken more than an extra page in the journal. In fact, it probably would not have taken even this amount of space if the lists in the article were consolidated and asterisks used to mark competencies considered inappropriate for the course or poorly attained by students.

FACTORS AFFECTING INSTRUCTION

Schafer, John E. "Effect of Individualized Goal-Setting on College Biology Students' Locus of Control." Journal of Research in Science Teaching 18(5):397-402, 1981.

Descriptors--*Biology; College Science; College Students; Higher Education; *Individualized Instruction; *Locus of Control; *Pretesting; Science Curriculum; Science Education; *Science Instruction

Expanded abstract and analysis prepared especially for I.S.E. by Lawrence L. Gabel, The Ohio State University

Purpose

The purpose of this study was to determine if shifts in students toward greater internality in locus of control could be produced by one, or both, of two means:

1. modifying a form of individualized instruction (audio-tutorial) to give more control to students for their own learning; and
2. repeated measures of locus of control.

Rationale

The investigator inferred from previous research findings that a student's perception of control of the environment is associated with academic achievement and that selected instructional methods could move students toward greater internality in locus of control. For the study, external locus of control was defined as a perception that reinforcement is the result of luck, chance, fate, control of powerful others, or is unpredictable because of the great complexity of surrounding forces. Internal locus of control was defined as a perception that a reinforcement is contingent upon personal behavior or relatively permanent personal characteristics (Rotter, 1966). Regarding locus of control, Bar-Tal and Bar-Zohor had reported previously (1977) that in their review of 36 studies they had found 31 of the studies reporting internal locus of control to be positively and significantly associated with academic achievement. A decade earlier, Coleman, et al. (1966) had reported students' sense of control of the environment, as well as other variables, to be related to students' achivements in grades five, nine, and twelve.

The investigator cited several studies which, when taken together, offered conflicting evidence that individualized instruction would promote students' shifts toward an internal locus of control. He summarized by posing general questions as the basis of his research. Can some additional modifications of a current form of individualized instruction promote shifts toward greater internality in locus of control? And, what is the effect of pretesting on posttest measures of locus of control?

Research Design and Procedure

The study population was all students enrolled during Fall, 1978, in an introductory zoology course at Purdue University. Having 40 divisions meeting some time between 7:30 a.m. and 4:30 p.m. four days a week, the course was taught via an audio-tutorial approach. Since students were scheduled into divisions by computer and since Herron and Luce (1978) found no systematic

bias in such scheduling techniques, the investigator secured the study's sample by randomly stratifying on divisional time. Thus, all students in divisions meeting at 9:30 a.m. became the study's subjects; these represented 8 percent of the study's population. These students were randomly assigned to the various groups of the research design.

Control group students studied by means of a "classic" audio-tutorial approach (Classic A-T). (Postlethwait et al., 1972) This consisted of a combination of independent study, small assembly, and general assembly sessions. Experimental group students differed from the control students in that they "used the Individualized Goal-Setting (IGS A-T) approach to Optional Minicourse mastery rather than the Classic A-T approach." Specifically, this included the components of:

1. student selection of minicourse goals and objectives;
2. student selection of instructional activities (minicourse, books, journal articles, demonstrations, instructors);
3. written quizzes with immediate question-by-question feedback; and
4. personal interaction with an instructor

The study's design was Solomon Four-Square (Campbell and Stanley, 1963) with two levels of treatment (Classic A-T, IGS A-T) and two levels of pretesting (pretest, no pretest). Data were collected using the Rotter Internal-External Locus of Control Scale (test-retest reliability of 0.89 and internal consistency of KR_{20} = 0.70) and the Academic Internal-External Locus of Control Scale (test-retest reliability of 0.84 and internal consistency of Cronbach = 0.86).

Findings

Efficacy of random assignment of subjects was checked using a multinomial chi-square test of cells. No significant differences among groups were found regarding sex or academic major. The two pretested groups were found to have no significant differences on selected variables of locus of control, cognitive preference, or course interest. Making these tests at the 0.10 probability level, group equivalence was concluded.

Converting the research purposes to research hypotheses and testing the null form by multivariate and univariate analysis of variance with unequal cell sizes, none was rejected at a 0.10 probability of committing a Type I error. Specifically these hypotheses were:

1. The mean posttest Rotter I-E and Academic I-E socres of IGS A-T students will equal the respective mean scores of the Classic A-T students.
2. The mean posttest Rotter I-E and Academic I-E scores of pretested students will equal the respective scores of unpretested students.

Interpretations

The investigator concluded that, given the conditions of the study, it was clear that Individualized Goal-Setting would not promote shifts toward an internal locus of control. Four points were made regarding this conclusion. First, the IGS A-T approach lacked potency since it was possible to extend IGS

A-T only to Optional Minicourses of the overall course. It was explained that Optional Minicourses contributed less than 10 percent to the course grade for three-credit students and less than 33 percent for four-credit students. Second, opportunities for goal-setting made available by IGS A-T were not utilized by students; that is, small percentages of students actually exercised their options to select or write unique objectives. Third, the investigator stated that there may be difficulties with predictive validity of locus of control in many classroom settings and that it is not that the students feel internally controlled but rather that they feel the situation is realistic, feasible, and capable of mastery. Fourth, it was recognized that selection of only 8 percent of the study population introduced a sizeable sampling error and, more importantly, that the small sample size made it difficult to obtain statistical significance.

ABSTRACTOR'S ANALYSIS

Two literature citings by the investigator which were of major consequence, Coleman et al., 1966, and Bar-Tal and Bar-Zohor, 1977, suggested strongly that internal locus of control is positively and significantly associated with academic achievement. Literature cited regarding whether or not individualized instruction or independent study could cause shifts in students toward internality of locus of control was conflicting although the majority of the citings by the investigator suggested that it could not. In fact, a summarizing statement was "...at a statistically significant level, there is little evidence to suggest that individualized learning conditions promote shifts toward an internal locus of control." A computerized search of the ERIC data base by this reviewer identified two additional studies in which the results supported the investigator's summarizing statement.

Given the conflicting evidence related to individualized instruction promoting shifts in students toward greater internal locus of control, the preponderance of evidence indicating that it does not, and the investigator's conclusion that little evidence existed to support the cause and effect relationship, this reviewer would mildly question the value in expending effort and resources to conduct the study. However, given that the investigator sought to determine if additional modifications of a form of individualized instruction, audio-tutorial, could cause the desired effect on locus of control, then the reviewer's questioning is somewhat mediated. The investigator's intent is likened to the necessity of a critical mass for self-sustaining fission of nuclei or to a required threshold level for a drug to be medically therapeutic.

With this intent in mind, this reviewer would like to have seen attention given to it in the introductory section of the report. Additionally, since one of the study's purposes was to investigate the effect of locus of control pretesting on end-of-course measurement of locus of control, this reviewer would have appreciated some attention to the purpose in the introductory section of the report; none was given.

Although the study made no new methodological contributions, it did lend support to Herron and Luce's (1966) finding that systematic bias is not produced when students are assigned by computer to large, multi-divisional courses. The investigator performed appropriate checks and described well the results of the statistical checks regarding the process to randomly secure and assign subjects to the various groups.

This reviewer believes the Solomon Four-Group Design was a variable design to use to investigate simultaneously the study's two purposes. It is depicted as:

R	O_1	X	O_2	where R = Randomization
R	O_3		O_4	X = Treatment
R		X	O_5	O = Measurement
R			O_6	

Campbell and Stanley (1963) claim the design controls for all usual sources of internal invalidity and one of three usual sources of external invalidity. Interaction of selection and treatment, although its control is questionable by the design, was shown not to be a threat to external validity; this was discussed in the preceding paragraph. The remaining potential threat to external validity, over which this design has control, was not discussed by the investigator. This threat is related to subjects of an experiment being inadvertently exposed during the course of the experiment to the experiment's treatment but in another non-experimental setting.

Regarding the adequacy of the written report, this reviewer was led to believe from a reading of the descriptions of the different instructional systems administered to the treatment and control groups that they were uniquely different. However, upon reading the discussion section, and in addition the abstract, it became clear that such was not the case. This led to questions as to why the methods sections were not more clearly written. Was it unintentional? Was it that the treatment was perceived before the study was undertaken that it would be unable to produce the hypothesized result? Why was a pilot study not undertaken prior to the major study given the explicitness of the explanations in the discussion section as to why the anticipated effect on locus of control was not realized?

It was stated previously that the actual purposes for which this study was undertaken were addressed indirectly in the introductory section of the report and in the case of one of the purposes, not at all. This latter purpose, related to effects of repeated measures on later measures of locus of control, again was not discussed at all in the discussion section of the report. Why in a study having two purposes, and two respective hypotheses for which a strong research design was used, would one of the purposes not be addressed in the introductory or the discussion sections of the report?

This study is another example of the many which report no significant findings. The question must always be asked regarding the outcomes of these studies as to whether the studies were conceptualized and conducted in such a manner that truly the findings are valid that the null hypotheses could not be rejected. In this particular study, probably enough evidence has been presented as to suggest a response to this rhetorical question. If not, the final evidence is provided by the investigator's statement in the discussion section that the study contained a sizeable sample error and his recognition that small sample sizes make it more difficult to obtain statistical significance.

REFERENCES

Bar-Tal, D. and Bar-Zohor, Y. "The Relationship Between Perception of Locus of Control and Achievement: Review and Some Educational Implications." Contemporary Educational Psychology, 2:181-199, 1977.

Campbell, D. and Stanley, J. Experimental and Quasi-Experimental Designs for Research. Chicago: Rand McNally College Publishing Company, 1963.

Coleman, J., et al. Equality of Educational Opportunity. Washington: U. S. Government Printing Office, 1966.

Herron, J. D. and T. Luce. "Return to 'the Proper Experimental Unit'." Journal of Research in Science Teaching, 15:293-302, 1978.

Postlethwait, et al. The Audio-tutorial Approach to Learning, Third Edition, Minneapolis: Burgess Publishing Company, 1972.

Rotter, J. "Generalized Expectancies for Internal Versus External Control of Reinforcement." Psychological Monographs, 80, (Whole No. 609), 1966.

Lynch, P. P. "Scientific Language and the High School Pupil: Preferential Thinking Styles." Journal of Research in Science Teaching, 18 (2), 131-138, 1981.

Descriptors--Cognitive Development; *Cognitive Measurement; *Cognitive Processes; *Concept Formation; *Concept Teaching; Generalization; Science Education; *Scientific Concepts; Secondary Education; *Secondary School Science

Expanded abstract and analysis prepared especially for I.S.E. by Richard M. Tolman, Brigham Young University.

Purpose

The purpose of the study was to investigate the effect of the independent variable: grade level in Australian high schools on the dependent variable, preferential style of thinking. Three levels were specified for the dependent variable: membership, partial association, and generalization.

Rationale

The study was based on a previous paper (Lynch et al., 1979) reporting the result of a study of student responses to a multiple choice test of 16 definitions of physical science terms. The 16 terms are atom, gas, mass, neutron, molecule, volume, proton, solid, compound, length, element(s), liquid, area, mixture, electron, and ion.

In the paper being abstracted, the 16 terms from the previous examination were used again. However, instead of listing one correct answer and three distractors, the author included four possible correct responses (definitions) for each term. One definition was listed as a membership (M) in which the concept is understood in terms of a single or group of easily recognizable object(s); e.g., H represents an atom of the element hydrogen. Two definitions were listed as two different partial associations (P), in which the concept is defined in terms of a specific characteristic or attribute; e.g., about 100 different kinds of atom exist. An atom is a neutral particle. One definition was listed as a textbook generalization (G) in which the concept is defined in terms of a generalization involving a frame of reference considered to be particularly appropriate; e.g., an atom is a small particle of matter.

A major unstated assumption is that the students' selection of a "most preferred" response will indicate the type of thinking that led to the selection. This assumption will be addressed in the analysis section of the abstract.

Research Design and Procedure

The research design used in the study does not fit any particular design discussed by Campbell and Stanley (1973). The closest design cited by Campbell and Stanley would be the pre-experimental, one-shot case study. The author, however, used four groups instead of one in a quasi time-series design. Responses were collected from students near the end of the seventh, eighth, ninth, and tenth grades in an attempt to study changes in thinking styles over time. The design could be expressed as follows:

$$\begin{array}{ll} X & O_1 \\ X & O_2 \\ X & O_3 \\ X & O_4 \end{array}$$

Even though the experimental variable of event, X, was not stated explicitly, one must assume that this would be exposure to the Tasmanian educational system. The O's represent the measurement obtained for each group (grade levels seven through ten) with the 16-item sentence preference test.

The author reports that 1635 students were selected, nonrandomly, from six Tasmanian high schools (students from 12 to 16 years of age). The high schools were chosen as being representative of the high school population in Tasmania, taking into account, "catchment area, urban and rural environment, and potential socioeconomic status of parents". The total sample consisted of 969 boys and 666 girls. The imbalance in number of boys was due to the inclusion of a boy's private school. The other five schools are coeducational. The numbers represented at each grade level were seventh-481; eighth-504; ninth-384; and tenth-241.

The pupils were informed that all four of the sentences for each term in the test were correct and that the students should read carefully through all of the sentences. They were instructed to pick out the sentence that they preferred as having the most meaning for them when they thought about the underlined word. Pupils then indicated their choice by placing a mark next to the appropriate sentence.

The percentage of responses was calculated for each of the categories (M, P, and G) for each of the 16 items. This was completed for each of the four grade levels. The average percentage response for M, P, and G was also calculated across all 16 questions for each of the four grade levels and displayed in Table II. A z test was performed for each type of response (M, P, or G) for each of the 16 words. The percentage chosen for grades seven and ten were subjected to the statistical analysis. This resulted in a total of 48 z tests-three for each of the 16 words. An asterisk was used to indicate where a minimum significance level ($P < .05$) had been reached for differences between seventh and tenth grades.

IQ estimates were also obtained for students in the sample. Correlations were computed between IQ and sentence preference types.

Percentages of M, P, and G responses for the total test were displayed and plotted against grade level in Fig. 1. No tests of significance were included with these data. Table III shows the direction of change, across all four grade levels, for each preference type within each word being defined. The most preferred type for each word was listed for grade seven and grade ten. Arrows were included to illustrate the direction of the change. No tests of significance were included with this table.

Findings

Membership preference from grade seven to grade ten, where a change exists, decreases. Only three terms, "gas", "liquid", and "solid" received a membership score higher than ten percent. The membership level at grade ten for "gas" is

twenty percent. Both generalization and partial association are increasingly preferred as grade level increases. The one exception to this finding is for the term "atom", where there is also an increase in preference for membership. The terms "gas", "liquid", and "solid" are preferred at the level of partial association from grade seven through grade ten. The terms "electron", "neutron", "length", and "volume" show a transition in preference from partial association to generalization over the period of time.

The average response level for membership, when calculated across all sixteen terms, decreases slightly from grade seven through grade ten. The same result occurred for the response level for partial association. The average response level for generalization increased from grade seven through grade ten.

Correlation coefficients (for n = 548) were computed between IQ estimates and levels of preference. The coefficients are as follows: M pref. (-.16); P pref. (.18); G pref. (.36).

Reliability estimates (KR-20) were reported for each level of preference. These are membership, .46; partial association, .68; and generalization, .65.

Interpretations

The consistency of trends for individual items; the interrelationship of the M, P, and G total scores with grade level; the correlations with IQ estimate; and the reliability of the test were cited as strong evidence for the validity of the phenomenon of preferential thinking style, as it applies to the pool of sixteen science terms used in the sentence preference test.

Two general trends were cited as having implications for the teaching of science. First was the "marked apparent transition of preferences in Grade 7". Second was the "...ceiling in preference categories that is reached by the end of Grade 10". Grade 7 is cited as representing "...a nexus of change in pupils' thinking style that needs to be recognized by teachers".

The maximum performance level obtained for the generalization category, "...implies limits for understanding. This particular level of thinking would seem to be a prerequisite for a full understanding of the appropriate high school texts".

ABSTRACTOR'S ANALYSIS

Relationships of the Study to Other Studies in the Area of Research

The research reported in this paper is casually related to several fields of research in cognitive reasoning abilities of secondary level students. The author's previously published work (Lynch et al., 1979) obviously established the germinal ideas for the study. The abstractor could not locate any similar studies via a computerized search of the ERIC data base.

The preferential thinking styles cited by the author apparently originated with him. The author did not cite any references for the thinking styles he reported. Different levels of meaning for adolescent learners, with reference to nonsubject specific sentence preference tests, have been discussed by Peel (1975). The levels of meaning, as used by the author, were conceptualized by him.

Four specific related areas of research come to mind. First, information processing research in cognitive psychology is a very fruitful area at the present time. Larkin's (1977a, 1977b, 1978, 1980, 1981) work in physics problem solving is most closely related to Lynch's work. The one-on-one interview techniques and the structure of the students' problem solving strategies, as employed in information processing research, would provide valuable insights into the actual mental processes being used by students while responding to the sentence preference test.

Second, the research reported by Wollman and Lawson (1978), Renner and Phillips (1980), and others in research related to Piagetian type reasoning could be very helpful in this new area of research. Lynch cannot be faulted for not basing his research within the Piagetian paradigm, and such criticism is not intended. Rather, it is suggested that the relationship between the results of the sentence preference test and levels of Piagetian reasoning that are being used could provide some additional valuable insights into the research being initiated by Lynch.

Third, the Concept Structuring Analysis Technique (CONSAT), as developed by Champagne and Klopfer (1980), could also prove to be very helpful and possibly provide additional insights into the reasoning being utilized by students responding to sentence preference tests. One wonders what relationships the students could construct among the sixteen terms used by Lynch. The CONSAT technique could be a most useful tool to use in probing more deeply into students' thought processes.

Fourth, hemispheric specialization may have some influence on the results obtained by Lynch. Bogen (1975) provides an excellent review of hemispheric specialization and includes 95 sources in his paper. This area of research could shed some valuable insights into Lynch's results.

New Conceptual Contributions of the Study

The categorization of students' sentence preferences into membership, partial association, and generalization types appears to be unique to this study. These definitions could be used in appropriately designed follow-up studies to Lynch's report.

Validity of the Study

A serious question needs to be answered at this point. The application of a sentence preference test and the categories of preference are certainly valid. The validity question being raised here is whether or not the students' responses really represent preferential thinking styles or merely a preference for a particular structure in a written definition. Also, could previous exposure to sentence structure be exerting an influence? Whether or not any particular thinking style is being utilized could be verified through in-depth interviews with a random sample of students who expressed a preference for one of the sentence types used. The validity of a paper and pencil test, such as the sentence preference test, to uncover actual preferential thinking styles is doubtful. Any one or a combination of the three related areas of study cited previously could help resolve this validity question. Additional research must be conducted to validate this procedure for representing thinking styles.

It would be well for Lynch and his colleagues to conduct further validation studies before attaching any significance to the results of the study.

The validity of one of Lynch's conclusions is also questionable. He reports that, "There would seem to be a crossover of generalization and partial association somewhere in the early part of Grade 7.... Consequently, Grade 7 is seen to be an interesting transitional period for high school pupils with regard to preferential thinking style."

Unless Lynch has gathered data on pre-seventh graders that were not reported in this study, it is inappropriate to conclude that a "crossover" has occurred. A conclusion such as this must be based on baseline data from students who are younger than the seventh graders reported in the study.

Reliability estimates should have been reported for the test as a whole, not for each particular response type. This application of the Kuder-Richardson formula is not valid.

Comments on the Research Design

Considering the preliminary nature of the research, the pre-experimental design is appropriate. True experimental designs would not have been appropriate at this stage of the investigation.

The statistics used are appropriate for the design, but they leave a reader wondering why they were used. The seventh- and tenth-grade proportions were entered into calculations for the z test and significant differences were reported. Several other alternatives seem reasonable.

First, the raw frequencies for each cell of the design were tallied and used to compute the reported percentages. The raw frequencies for all four grades could have been used in a chi-square test for each of the sixteen items in the sentence preference test. The resulting chi-square statistic would test for association between the levels of preference (M, P, or G) and grade level. More important, however, is the fact that data from all four years of the study would have been entered into the analysis, not just data from the first and last years.

Three examples from Lynch's data will illustrate this point. For the term "atom", the percentage of preference for membership for grades seven, eight, nine, and ten are, respectively, 7.8, 6.3, 4.9 and 8.5. Lynch's analysis indicates an increase from 7.8 to 8.5, but this ignores the decrease that occurred in grades eight and nine.

A different type of result is seen for the term "gas." The membership preference percentages for grades seven, eight, nine, and ten are, respectively, 23.5, 18.0, 15.8, and 20.5. Lynch's analysis indicates a decrease of three percentage points from seventh to tenth grade. This ignores the drop to 18.0 and 15.8 for grades eight and nine, followed by an increase to 20.5 for grade ten.

One last example for still another trend for the term "mass." Lynch's analysis utilizes a small drop in the percentage of preference for membership from seventh to tenth grade, with the respective figures being 5.9 and 5.4. However, the data for grades eight and nine indicate an increase to 7.6 for both years, followed by a decrease to 5.4 at grade ten.

Similar results occurred for other terms, but the three examples cited serve to illustrate the point. Entering all of the data could influence the conclusions of the study.

A second possibility would be to utilize a time-series analysis. Depending upon the statistical program utilized, this could necessitate a random reduction in sample size at some grade levels to provide equal numbers for each grade level. The time series or trend analysis routines seem much more appropriate given the fluctuations of the data in grades eight and nine.

Comments on the Written Report

There are several difficulties with the written report that cause some problems for the reader. The first problem relates to the description of the sample used in the study. Lynch reports that 1635 pupils comprised the sample. The 969 boys and 666 girls add up to and verify the cited total numbers. However, the numbers cited for each grade level do not add up to 1635. The grade-by-grade numbers are as follows: seventh-481; eighth-504; ninth-384; tenth-241. Summing across these figures yields a total of 1610 pupils, a discrepancy of 25.

A second problem relates to the description of the samples. Lynch reports that, "Grades 7 and 8 are total samples". One must assume that this means all of the pupils in grades 7 and 8 from the six schools are included in the sample. This could merely be a semantics problem, but one additional sentence stating the exact situation would have helped. The major problem is with the ninth and tenth grade samples, when Lynch reports, "Grades 9 and 10 are representative samples of approximately 2 class equivalents from each school". Here, the question arises about whether two classes were selected from each school or merely enough students to comprise two classes. More importantly, the reader desires to know how these students were selected for inclusion in the sample. How many students were omitted? Addressing these questions and clarifying the exact status of the subcomponents of the sample would have helped the accuracy of the report.

A third problem occurs under the heading "Application of the Test". Three sentences were included under this heading as follows:

"The sentence preference test was given to all the pupils who, one month previously, had completed the recognition of concept definition test. Those who answered the earlier test correctly were treated as a single group. The data presented (Lynch et al., 1979) refers to this particular group."

The above quoted sentences reflect a common problem that occurs with many people when they write reports. They know exactly what they want to say and it is clear to them. However, to the reader who is encountering Lynch's study for the first time, a problem of intended meaning occurs.

Does Lynch mean that all 1635 pupils in this study also took part in the previous study? Is that how the ninth and tenth grade numbers were reduced? His statement that, "Those who answered the earlier test correctly were treated as a single group" also leaves many questions unanswered. What comprises, "answering the earlier test correctly"? How were these students treated as a single group? What were the criteria (100 percent correct?) for answering the test correctly?

This short section in Lynch's report needs considerable amplification to answer the questions that it raises. There appears to be a confusion about the constituency of the samples in the two studies. This type of problem could be overcome if colleagues who are unfamiliar with a study could be invited to critique a paper while it is still in the draft stage. Persons familiar with an ongoing study are, many times, too close to the data to render unbiased opinions and critiques.

Assessment of the Current State of Research in the Area of the Study

Research related to the categories of preferences utilized by Lynch is in its infancy. Considerable attention must be paid to the cognitive thought processes utilized by students while responding to the sentence preference test. Lynch's ideas are worthy of additional study by other investigators.

Suggestions for Future Research Direction and Effort

Four areas of related study have been suggested in this abstract as possible means of delving more deeply into the thought processes of students. Concept Structuring Analysis technique, information processing studies, Piagetian-based reasoning studies, and hemispheric specialization studies all seem appropriate methodologies for pursuing the research. All four approaches could provide valuable data for further studies.

REFERENCES

Bogen, J. E. "Some Educational Aspects of Hemispheric Specialization." UCLA Educator, 17:24-32, 1975.

Campbell, D. T. and J. C. Stanley. Experimental and Quasi-experimental Designs for Research. Chicago: Rand McNally, 1973.

Champagne, A. B. and L. E. Klopfer. "Using the CONSAT: a Memo to Teachers." Learning Research and Development Center, University of Pittsburgh, 1980.

Larkin, J. H. "Problem Solving in Physics." Working paper, Group in Science and Mathematics Education, University of California, Berkeley, 1977(a).

Larkin, J. H. "Skilled Problem Solving in Physics: A Hierarchical Planning Model." Working paper, Group in Science and Mathematics Education, University of California, Berkeley, 1977(b).

Larkin, J. H. "Information Processing Models and Science Instruction." Paper presented at the AAAS Symposium on Models of Learning and Their Implications for Science Instruction, Washington, D. C., February, 1978.

Larkin, J. H. "Understanding, Problem Representations, and Skill in Physics." Paper presented at NIE-LRDC Conference on Thinking and Learning Skills, Pittsburgh, October, 1980.

Larkin, J. H. "The Cognition of Learning Physics." In Press, American Journal of Physics, 1981.

Lynch, P. P. et al. "Scientific Language and the High School Pupil." Journal of Research in Science Teaching, 16:351-357, 1979.

Peel, E. A. "Predilection for Generalizing and Abstracting." British Journal of Educational Psychology, 45:177-188, 1975.

Renner, J. W. and D. G. Phillips. "Piaget's Developmental Model: A Basis for Research in Science Education." School Science and Mathematics, 80:193-198, 1980.

Wollman, W. T. and A. E. Lawson. "The Influence of Instruction on Proportional Reasoning in Seventh Graders. "Journal of Research in Science Teaching, 15:227-232, 1978.

Campbell, Richard L. "Intellectual Development, Achievement, and Self-Concept of Elementary Minority School Children," School Science and Mathematics 81(3):200-204, March 1981.

Descriptors--*Academic Achievement; Correlation; Elementary Education; Elementary School Science; *Intellectual Development; *Minority Group Children; Minority Groups; Science Education; *Self Concept

Expanded abstract and analysis prepared especially for ISE by Steve Tipps, University of Virginia.

Purpose

Two questions were investigated in the study. The first asked if the relationship cited elsewhere in the literature between self-concept and school achievement could be reaffirmed with the elementary minority students in the study. Second, would there be a relationship between the self-concept of the subjects and their cognitive development?

Rationale

Humanistic psychology which suggests self perception as the basis for learning serves as the theoretical basis for the study. The author cited numerous studies showing a relationship between self concept and school achievement. In addition, research on Piagetian intellectual development and its strong relationship to academic achievement was presented.

Research Design and Procedure

The study is descriptive in nature. Three measures were taken on a group of 51 students from a minority, lower SES, urban school. The school population was about half Hispanic and half Black, although the sample composition is not specified. The subjects were second graders with mean age of 8. The Stanford Achievement Test, Piers Harris Self-Concepts Scale, and seven Piaget conservation tasks were administered. Achievement scores from each of the subtests were correlated with the self concept scores, as were each the conservation tasks. The Piagetian tasks represented a scale from preoperational to formal operational thinking. Two scores within each conservation task were used in analysis--answer and explanation.

Findings

The correlation coefficients between achievement subtests in reading, social studies, science, and listening comprehension were significant and in the .40-.60 range. Five conservation scores had significant correlation coefficients although they were slightly lower than the achievement scores (.30-.40). The tasks positively correlated were the four higher conservation ones: length, area, weight, volume. Both answer and explanation scores were significant for length, while only explanation scores were significant for the other tasks.

Interpretations

The results confirmed the relationship between self concept and achievement in minority students. In addition, the positive relationship between self concept

and four higher Piagetian tasks indicated a need for enhancement of self concept before children can move to higher level thought. The author suggested that teachers and curriculum specialists should emphasize successful experiences building self concepts to increase both achievement and intellectual development.

ABSTRACTOR'S ANALYSIS

The new aspect of this study is examination of self concept and intellectual development. Although relationship of achievement with self concept and with intellectual development is well documented, whether there exists a separate relationship between self-concept and intellectual development has not been dealt with in the literature. The study may have begun to ask the question, but several issues must be raised about the method and results. Scoring of the conservation tasks is a particular concern. Dividing the score for Piagetian tasks into separate scores for answer and explanation is not standard. The task is usually considered a unity. No rationale for this scoring procedure is presented.

The answer, being "yes" or "no," may be no more than random. Use of Pearson's product-moment correlation might also be questioned due to the dichotomous nature of the scores. Combining the two scores would have been more consistent with previous research and might have changed the results. Understanding of the relationship between concept and conservation ability might also have been enhanced by correlating a total score on conservation with self concept. Interpretations of the study imply more than a correlation between the variables. The author may believe that self concept enhancement is a prerequisite to academic achievement and intellectual development, but neither this nor any of the studies cited demonstrate a cause-effect relationship. Without denying that motivation has a strong impact on learning or that success appears to influence self concept, the proposition that self concept determines success in learning or triggers intellectual development is not sustained.

A more dynamic interrelationship between achievement, self concept and intellectual development is possible. Investigating the complexity of these variables would call for the same or similar measures on larger populations over several age levels. Changing the scoring procedure for the Piagetian tasks and adopting a multivariate analysis, the author might make a considerable step in answering his research question.

Hudson, H.T. and Ray M. Rottman. "Correlation Between Performance in Physics and Prior Mathematics." Journal of Research in Science Teaching, 18(4):291-294, 1981.

Descriptors--*Academic Achievement, College Mathematics, *College Science, *Dropout Characteristics, Higher Education, *Mathematics, Mathematics Education, *Physics, *Predictor Variables, Science Education

Expanded abstract and analysis prepared especially for I.S.E. by Glen S. Aikenhead, University of Saskatchewan

Purpose

To test the hypothesis that mathematics ability influences both first year university physics grades and drop out rate from physics.

Rationale

Significant correlations between math ability and university physics achievement have been reported in the literature. However, because about 40 percent of science and engineering majors drop out of these career tracks by second year of university, the authors wonder what the parameters are which strongly affect the staying power of a student. The mathematics parameter was chosen on account of its prerequisite status to physics courses.

Research Design and Procedures

The investigation is a simple correlation study. A test, consisting of 28 questions (two per 14 topics in algebra and trigonometry pertinent to what is used in physics), was administered at the beginning of the year to 1403 students in an introductory, preprofessional noncalculus physics course. These students formed two groups, those who dropped out (490) and those who did not (913). The authors checked the content validity by consulting with "a number of instructors". No other type of validity was assessed. No reliability measures were obtained for the test.

Findings

The Pearson product-moment correlations between the math pretest and the final grade, a projected grade for the dropouts, were .418 and .232 for the completing and dropping students respectively (both are statistically significant at the p .001 level). The authors also allude to, but do not report systematically, the fact that credit in a mathematics course does not guarantee an anticipated achievement. Over 35 percent of the students who had obtained a prerequisite math credit scored less than 50 percent on the research math instrument.

Interpretations

The authors stated that the .418 correlation was consistent with previous studies. The authors interpreted the .232 correlation as being "weak" and, therefore, they concluded that math ability must have a secondary influence on drop-out status. Primary influences on drop-out status were stated as "yet unidentified" cognitive factors. The authors went on to discuss the fact that

the distribution of math scores for the two groups of students overlapped too much "to provide any hope of predicting completion or drop, except for the extreme ranges of mathematics scores". However, the authors believed (no data are reported) that the distribution of physics grades for the two groups did not overlap much, and thus the physics grades formed two distinct groups: "those who can do the operations necessary to the study of physics and those who cannot." Lastly, the authors generalized their findings to all other science courses.

ABSTRACTOR'S ANALYSIS

Physical science inquiry has achieved great success with its reductionist approach to investigations, where the researcher narrows the conceptualization of the problem to a workable small number of variables. In the present study the authors expressed concern over the large 40 percent drop out rate in first year physics. The authors addressed the problem in the reductionist tradition by narrowing in on math ability as a sole variable. Their approach logically followed the literature cited. But a general education problem (40 percent drop-out rate) has been redefined as a specific problem of math ability. There is a research literature in educational counselling which addresses the drop-out problem of first year university students. This larger perspective was ignored in order to do a study that spoke not to the general problem, but specifically and narrowly to those physics professors who thought that math ability was the prime factor in students dropping physics. The reductionist approach in this case would seem to limit the study's impact on the general problem of dropouts. It did eliminate math ability as a prime factor, if one believes the authors' unsupported contention that a .232 correlation coefficient is weak though statistically significant.

An interesting discussion occurred over (a) erroneously inferring math ability from the fact that a student has completed a math course on the same topics, and (b) the extensive overlap of math ability of dropouts and completers. The data prerequisite to these discussions were not reported however.

The authors' apparent lack of concern for validity and reliability of their assessment instrument reflects the context in which the study was done. The context is narrowly defined by physics professors who find credence in the hypothesis that math ability is a prime factor in drop outs. This readership appears to put a high value on content validity and a low value on measurement reliability.

Future researchers who are concerned with the general problem of drop-outs may discover significant variables by interviewing a random sample of drop-outs. Once the variables have been established, then variates can be chosen and quantitative studies could proceed. This empirical approach may be contrasted to the strictly logical approach seen in the present study where math ability was reasoned to be important, and where "other cognitive factors" were concluded to be most influential.

IN RESPONSE TO THE ANALYSIS OF

Hudson, H. T. and R. M. Rottman. "Correlation between Performance in Physics and Prior Mathematics." by Glen Aikenhead. Investigations in Science Education 9(3):58-59, 1983.

H. T. Hudson
University of Houston

An "abstract" should be accurate. A "critique" should be free of personal bias. Neither of these criteria was a burden assumed by the abstractor. The following points are pertinent:

1) Under "Research Design and Procedures," the abstractor indicates a concern that will later be reiterated as a criticism. This has to do with the method used to determine content validity. In fact, the procedure described in the paper paralleled exactly the example given on page 457 of the text, Foundations of Educational Psychology, by Fred Kerlinger. The abstractor might want to refer to page 459 of this text as to the general guidelines for content validity. For our purposes, the content validity was the most important criterion. The fact that the test scores are highly reproducible between classes over the years and that the Kuder-Richardson reliability has always been between 0.81 and 0.92 for all the classes given the test (before, during and after the study reported in the paper) is not the point. We didn't go into this because we felt it distracted from the thrust of the paper.

2) The abstractor ignores information given in the article and pretends this information is absent. ". . .the authors allude to, but do not report systematically, the fact that credit in a mathematics course does not guarantee anticipated achievement." We state that the topics of the test are taken from algebra and trigonometry. We further state that 95 percent of the students have credit for these subjects. We don't allude to anything. We come right our and say that 35 percent of the students score less than 50 percent on the test.

3) Under the section called "Interpretations," the abstractor makes statements that suggest a lack of understanding of statistics:

a) ". . .the authors believe (no data are reported) that the distribution of physics grades for the two groups did not overlap much." (emphasis mine) In fact, the data are very much presented, under the label "Physics Grade" in Table 1. As given there, the mean physics grade for the completes was 56.7 with standard deviation 16.6 and the mean physics grade for the drops was 26.7 with standard deviation 18.6. Any "researcher" or "abstractor" who does not realize that these numbers do indeed represent two distinct groups needs to learn basic statistics.

b) A correlation coefficient of 0.232 explains only 5 percent of the variance. We said that is weak, and it is. Apparently the abstractor is unaware that larger sample size can artificially inflate "significance." On this point, one can refer to any good statistician.

4) In the "Abstractor's Analysis" section, the abstractor suggests that we should have used a series of research reports from education counselling which addresses the drop out problem of first year university students. We did not, for two reasons. The first reason is that it would be inappropriate to generalize drop out reasons for first year students to this sample. Only 11.3 percent were freshmen. The second reason is that we had drop out data from the population of students at the university where the study was conducted. On a "random sample" survey (see next paragraph) conducted by the university among students who had dropped out of school, it was found that students dropped out for many reasons, ranging from dissatisfaction with the parking to fear over personal safety. Since there was little overlap between the reasons students gave for leaving school and for leaving physics (again, see next paragraph), we concluded that very little of this information would be of use in our study. Our approach was a little more scientific.

5) There is a criticism that we didn't interview a random sample of the students who dropped. Actually, we interviewed <u>every</u> student who dropped over a two year period. By far the most frequently-given reason for dropping physics was that "I can't do the mathematics" or "I can't keep up," or a combination of the two, with speculation on the part of the student that maybe the two reasons were related. Since these students had credit for the prerequisites in mathematics, and many had "A" grades, we concluded that by their own perception their mathematics was inadequate. The original purpose of the math test was to demonstrate to the students just exactly what mathematical skills were expected of them. The consultation with a "number of instructors" was to guarantee content validity. When we told the students that the mathematics of the test would be necessary to work the problems of the physics course, we wanted to be certain that we were being honest. This study, which used that test, was to look at the impact of that mathematics as it impacts the performance in physics. That is what we did.

This abstractor has the right to criticize us for sins of omission, for not presenting all the facts. He/she even has the right to disagree philosophically with the approach used in the study. However, no one has the right to distort the facts and suggest that poor experimental technique would be better than rigor. As well, no one has the right to criticize any work simply on the grounds that there is philosophical disagreement, and to further make that criticism an innuendo.

Goldsmith, R. H. "Readability and Filmstrip Selection in Science." School Science and Mathematics 80(4):335-338, April 1980.

Descriptors--Difficulty Level; Educational Media; Educational Research; Educational Resources; *Elementary-Secondary Education; Elementary School Science; *Evaluation Criteria; Filmstrips; Instructional Materials; *Readability; Readability Formulas; *Science Education; Secondary School Science; *Selection.

Expanded abstract and analysis prepared especially for I.S.E. by John Edwards, James Cook University of North Queensland.

Purpose

To determine the readability of the captions from a wide range of middle and high school science filmstrips, and to compare the measured reading level with the level suggested by the publisher.

Rationale

The investigation is planned to add another dimension to the analysis of filmstrips. Readability is one criterion which has not been generally considered relevant to captioned filmstrips. The captions are run together and treated as continuous text for analysis purposes. Readability is measured using two semantic measures: the Dale-Chall Readability Formula, and the Fry Formula.

Research Design and Procedure

Seventy filmstrips were analyzed. They were taken from eight different middle and high schools in which they were used either in the media center or the classroom. Each had a level of usage recommended by the publisher. The filmstrips covered a range of publishers and publication dates. Topics covered included: botany, weather and space sciences, geology, biology, physics, chemistry, and health and medical sciences.

For comparison with measured reading levels, the filmstrips were categorized into four groups, based on the publisher's recommended usage level:

Group I - recommended usage level of 6-8 (middle school group).
Group II - half level 7-8 and half level 7-9 (junior high school).
Group III - half level 7-10 and half level 7-12, with one at 7-11 (junior-senior high schools).
Group IV - level 10-college, with one at 10-12.

Each filmstrip was analysed with both the Dale-Chall Readability Formula, and the Fry Formula.

Findings

The findings of the study are completely summarized in two tables as shown below:

Table 1--READABILITY LEVELS OF SEVENTY SCIENCE FILMSTRIPS BY GROUPS

Filmstrip Group	FORMULA Dale-Chall	Fry
I	9±3.5	7±3.7
II	11±2.7	9±2.6
III	10±2.0	9±2.1
IV	13±1.5	11±1.5

TABLE 2--READING FORMULAE AND READABILITY CALCULATIONS FOR ALL FILMSTRIPS

Formula Used	% Above Minimum Level*	Mean # Years Above Minimum Level *	% Outside Range*
Dale-Chall	80%	+4.3	51%
Fry	64%	+3.2	27%

* As suggested by publisher
Calculated only for those filmstrips above minimum level

As can be seen:

(a) The Dale-Chall results report average readability at the top of the publisher's indicated range, while Fry results match the midpoint of the publisher's range, or slightly above that point.

(b) Using the Dale-Chall, 51 percent of filmstrips had readability levels outside the publisher's range. The Fry Formula suggested 27 percent outside the publisher's range.

Interpretations

Goldsmith concludes that teachers ought not to rely on publisher's suggested usage levels for filmstrips, especially if independent study use is contemplated. Teachers should employ readability formulae in filmstrip selection and plan filmstrip usage in view of the findings. If the readability level is too high, the teacher can prepare the students beforehand, for example by introducing new vocabulary. Disparity between readability levels revealed by the two formulae is seen as a note of caution.

ABSTRACTOR'S ANALYSIS

Goldsmith's study is an interesting addition to the growing body of literature recommending focused analysis in materials development and materials selection. The findings support the concern felt by many practicing teachers about the effectiveness of captions in many filmstrips.

There are undoubtedly problems with the validity of applying the two readability formulae to passages made up from consolidated captions. Firstly, the design of many captions would not allow for valid consolidation. Secondly, captions are not meant to stand alone. They are closely integrated with, and supportive

of, the filmstrip. The use of two semantic measures of readability, while useful for cross-checking, reduces the power of the study. To a large extent some of these problems would be overcome by the use of a measure of functional readability, such as Cloze procedure. In this way the actual filmstrip could be shown, with appropriately modified captions, to sample students. This would be a much more time consuming study than that of Goldsmith. However, if the potential problem of filmstrip readability is to be tackled in depth, then such a study would be needed. Cloze procedure is a more practical measure for this purpose and would probably reveal a more depressing picture than that shown here. Of course, the main use for such an analysis ought to be in the design and trialling of filmstrip captions, rather than in revealing the strengths and weaknesses of finished products.

The use of means in reporting results in Table 1, while revealing general trends, is largely counter-productive. Means can also be helpful for concealing trends. A more revealing presentation of data would have been to give the actual readability breakdowns for each group with each formula. In Table 2, the "above minimum level" could be seen as a harsh judgment. Publishers could just as validly ask for a "below maximum level" figure. Again, a more detailed breakdown of the figures in the last three columns would provide a potentially more useful data source for the reader.

So, despite methodological weaknesses, Goldsmith's study serves the useful purpose of focusing attention on an area in need of further research. Filmstrip producers ought to, preferably voluntarily or in response to critical professional analysis, pay close attention to the readability of filmstrip captions for expected audiences. Such analysis needs to be part of the development process, rather than a means for revealing expensive failures.

ATTITUDES

Kiely-Brocato, Kathleen. "An Assessment of Visitor Attitudes Toward Resource Use and Management." Journal of Environmental Education, 11(4):29-36, Summer 1980.

Descriptors--*Administration; *Attitudes; Attitude Change; Attitude Measures; *Beliefs; Environmental Education; *Natural Resources; Outdoor Education; Parks; *Recreation; Questionnaires; *Surveys

Expanded abstract and analysis prepared especially for I.S.E. by Ronald D. Simpson, University of Georgia

Purpose

This investigation had two objectives. The first one was to assess visitor attitudes toward resource use and management at Shenandoah National Park. To accomplish this objective a method for scaling attitudes which alleviates some of the problems of interpreting the effects of two related components of attitude--beliefs and corresponding evaluations of those beliefs--was developed. The second objective was to formulate guidelines that suggest how to design an effective interpretive message taking into account visitor attitudes toward resource use and management.

Rationale

The investigator accepts the premise that attitude is comprised of two components: an affective or emotional component and a cognitive or belief component. In other words (according to Fishbein), an individual's attitude toward a given object is related to a belief that the object possesses certain attributes and feelings held by that person toward those attributes.

The investigator also accepts Fishbein's formulation that allows for considerable latitude in measuring attitude. In this study attitude scores are computed using a unipolar scale for beliefs and a bipolar scale for evaluation of those beliefs. The investigator believes that this scaling technique may be particularly suitable for assessing visitor attitudes toward resource management programs.

Research Design and Procedure

A questionnaire was developed by the investigator to assess visitor attitudes toward resource use and management. The first portion of this task involved asking respondents, via 44 belief statements, whether or not they thought each item represented a policy or action taken by park managers at Shenandoah National Park. Response categories included strongly agree, agree, slightly agree, don't know, slightly disagree, disagree, and strongly disagree. In an attempt to promote equal psychological intervals between responses, the investigator chose to numerically anchor the responses from +3 to -3.

The second portion of this measurement device involved assessing how participants felt toward each of the 44 belief statements. This was accomplished by constructing a modified Likert-type scale. This technique, the investigator points out, complies with Fishbein's requirement that attributes be evaluated for each belief measured. Again, response choices ranged from strongly favor thorugh neutral to strongly oppose and were numerically anchored from +3 to -3.

A stratified random sample of days was selected to contact park visitors at Shenandoah National Park. The selected days were randomly matched to various sites: lodge, backcountry, station, or campground on designated sampling days visitors were contacted as they registered at these locations. Names and addresses were obtained on an information release form and in September, 1978, individuals were mailed the attitude survey. A followup questionnaire mailing ensued. The overall response rate was 81 percent.

Findings

The first objective of this study was to assess visitor attitudes toward resource use and management in Shenandoah National Park. This was done using the scaling system developed within the framework of Fishbein's expectancy-value attitude theory.

Each attitude item score represented an attitude toward a specific management policy or action at this park. Item scores were computed by multiplying belief about the policy or action by evaluation of the attribute. This technique is based on Fishbein's theory and is represented algebraically as:

$$A_o = b_i e_i$$

where: A_o = the attitude toward some object o.

b_i = belief i about object o, the subjective probability is related to some attirbute i.

e_i = the evaluation of attribute i.

n = the number of beliefs.

i = an attribute (another object, concept, characteristic, goal, value, etc.) associated with object o.

An overall attitude score for each respondent was computed by summing the 44 attitude item scores. The mean overall score was 42.813 (SD = 17.490). The potential range was from -54.064 (maximally unfavorable) to 130.680 (maximally favorable). The investigator concluded that, although visitors possessed an overall favorable attitude toward resource use and management at Shenandoah, the degree of favorableness might be increased substantially.

The second objective of this study was to suggest how to design possible interpretive messages which might positively increase visitor attitudes toward resource use and management at Shenandoah National Park. For each attitude item, the four following response categories exist:

1. an individual can favor the action ($+e_i$) and believe it is an agency policy ($+b_i$);
2. an individual can oppose the action ($-e_i$) and believe it is not an agency policy ($-b_i$);
3. an individual can favor the action ($+e_i$) and believe it is an agency policy ($-b_i$); and
4. an individual can oppose the action ($-e_i$) and believe it is an agency policy ($+b_i$).

A positive attitude score arises for Cases 1 and 2, and a negative attitude score arises for Cases 3 and 4.

The investigator goes on to state that, within Fishbein's conceptual framework, an interpretive message will change attitude only when the message alters the sum of belief/evaluation cross-produces ($b_i e_i$). Using this formulation, designing a specific influence attempt for each attitude item representing a Park Service management policy or action would be easy to do, but not practical. Therefore, the 44 attitude items in this study were divided into three groups based on frequency distributions of attitude item scores for each policy or action.

Belief and evaluation crosstabulations for each attitude item were arranged using a quadrant analysis. The four possible belief (b_i)/evaluation (e_i) combinations were: $+b_i +e_i$, $-b_n -e_n$, $-b +e_n$, and $+b_i -e_i$. For each policy or action, the percentage of attitude item scores in each quadrant was computed. This allowed for an a posteriori division of attitude items into three groups. For each policy/action in the first group, most of the attitude item scores were located in the $+b_i +e_i$ quadrant. While less than half of the item scores were located in the $+b_i +e_i$ quadrant, fewer than 25 percent were in $-b_i -e_i$.

Interpretations

Attitudes toward resource use and management at the national park in this study were generally positive. The investigator states that "In order to improve overall attitude scores (A_o), it is necessary to increase attitude item scores (a_i). Each attitude item score is the product of belief about a specific resource management policy or action (b_i) and corresponding evaluation of that policy or action (e_i). Ideally, an interpretive message designed to increase visitor attitudes should state agency policy or action and explain the benefits or positive consequences of the practice. However, when many visitors already favorably evaluate a practice, simple statement of agency policy or action may result in an increase in the percentage of positive attitude scores. As the number of persons favoring a policy or action decreases, it becomes necessary to also explain the positive consequences or benefits of the practice. It is particularly important that interpretive messages attempt to influence both belief and evaluation when many people initially oppose an agency policy or action ($-e_i$), but do not believe it actually is a Park Service practice ($-b_i$)."

The method of analysis used in this study is a way of considering visitor attitudes before attempting to design programs for attitude change. Also, it is a method that allows priorities to be established when resources are limited. Finally, it is a method that allows for monitoring of attitude change. This monitoring can be used to detect differences over time, across seasons, and among different groups.

ABSTRACTOR'S ANALYSIS

This investigation sought to assess visitor attitudes toward selected management practices at a national park and to incorporate the responses into future policy-making at the park. The study, in my opinion, was well done and represents a novel way in which beliefs and feelings can be assessed and the results used to make future decisions.

The investigator's conceptualization of the construct attitude was based on work by Fishbein, particularly his work on an expectancy-value model defining the mathematical relationship between beliefs about an object and attitudes toward the psychological object. Fishbein's work is highly respected in the social and behavioral sciences; hence, the investigator built her study on firm theoretical ground.

This study is tied to past scholarly work on attitudes in that it is based on the premise that attitude is comprised of both cognitive (belief) and affective (feelings) components. Again, I consider this premise to be a solid one, one that would be supported by most social and educational psychologists. The novel part of this study, the part that offers both conceptual and methodological contributions, is the scheme applied to measure both what people think exists (belief) and how they feel about what exists--or does not exist. In this study the investigator used Fishbein's model to assess whether or not visitors thought certain policies such as "prohibiting touching of animals" and "not controlling insects that attack trees" actually existed or not, and to assess whether or not visitors agreed or disagreed with each policy. To me, this study elucidates how one can measure these two components of attitude (belief and feeling) and use this information as feedback and input in an organization. The major contribution of this study to the field of attitude research, therefore, is the application of a model which allows one to assess beliefs and at the same time evaluate those beliefs. The following figure is taken from the Keily-Brocato report and represents a matrix of scores based on her extended version of the Fishbein Model.

		BELIEF						
		Strongly Agree	Agree	Slightly Agree	Don't Know	Slightly Disagree	Disagree	Strongly Disagree
		.990	.794	.598	.5	.402	.206	.010
EVALUATION								
Strongly Favor	+3	+2.97	+2.382	+1.794	0	−.030	−.618	−1.206
Favor	+2	+1.98	+1.588	+1.196	0	−.020	−.412	−.804
Slightly Favor	+1	+.990	+.794	+.598	0	−.010	−.206	−.402
Neutral	0	0	0	0	0	0	0	0
Slightly Oppose	−1	−.990	−.794	−.598	0	+.010	+.206	+.402
Oppose	−2	−1.98	−1.588	−1.196	0	+.020	+.412	+.804
Strongly Oppose	−3	−2.97	−2.382	−1.794	0	+.030	+.618	+1.206

Figure 2.--A Proposed Matrix of Attitude Scores Based upon an Extended Version of Fishbein's Model

Using this model, the investigator converted the belief scale, traditionally established as a bipolar scale, to a unipolar scale. Three reasons cited in her paper are:

> "First, the contribution of belief and evaluation to attitude were equally weighted prior to recoding since both were scaled in the same manner. In other words, a +3 on belief was implicitly equal to a +3 on the corresponding evaluative dimension. Although there are many definitions of attitude, most seem to stress the concept of evaluative reactions based on cognitive processes (10, 11). Thus, the evaluative dimension is of prime importance to the concept of attitude.

"Second, attitude scores derived from the use of the standard scaling procedure would not be managerially relevant. Managerially different scores are needed to assess visitor attitudes towards policies and actions. The most positive attitude scores should be obtained when a person favors an action and believes it is being done ($+b_i+e_i$). A positive attitude also should result when a person opposes an action, but doesn't believe it is being done ($-b_i-e_i$). A negative attitude should be computed when an action is favored, but the individual doesn't believe it is being done ($-b_i+e_i$). The most negative attitude should result when an action is opposed and the person believes it is being done ($+b_i-e_i$). This logical ordering of the four possible cases is not accomplished until the bipolar belief scale is changed.

"Finally, data interpretation would be ambiguous. For example, if a manager were interested in response to a specific policy, he could not ascertain whether numerous high attitude item scores were a result of many people strongly opposing the action (-3) and strongly believing it was not the agency policy (-3). Both give an attitude item score of +9."

The investigator also recoded the equal interval six point scale by anchoring it at .990 (strongly agree) and .010 (strongly disagree) and not including the zero point (don't know). In doing this the author very cleverly eliminated the zero points that exist on traditional scales, yet stayed within Fishbein's theoretical framework. In fact, Fishbein and Ajzen state that only beliefs which stand out above others (they refer to these as salient beliefs) contribute substantially to an attitude. The "don't know" response implies an unawareness to the policy or issue. Such a response, according to the investigator, represents a nonsalient belief and should not contribute to the attitude of interest.

By recoding both the belief and evaluation components, the investigator was able to weigh more heavily the evaluation factors. Also, she was able to order logically the four belief-evaluation quadrants so that numerical values within these cells were consistent with the theoretical framework within which she was working.

I found the proposed matrix by this investigation very sound and adaptive without violating the original version of Fishbein's Model. The investigator was able to, through this extended version, propose a scheme for measuring beliefs and feelings that were more appropriate for the needs of this investigation. This proposed adaptation coupled with display of how this model can be used to measure both beliefs and feelings (evaluations) is the major contribution of this study.

In addition to the unfolding and application of this model, I found the overall quality of this research report very high. The design used appeared appropriate and the manner in which it was presented was thorough, yet easy to follow. The author was also careful to document important sources of information such as the theoretical work of Fishbein, Fishbein and Ajzen and Nunnally. In short, this work was founded on sound theoretical underpinnings.

The only weakness I see in this study is associated with the survey instrument. This paper does not report the basis on which items were selected and reviewed. Reliability estimates are not given and no attempts are made to develop a case for validity relative to the 44 item scale. If important policy decisions are to hinge on results from questionnaires or similar attitude assessment instruments, we need to be sure that the instruments are valid and reliable. While validity is an attribute that usually takes longer to build, it is appropriate to address the issue with what information is available. In short, if the policies and practices of an institution are to be shaped and altered by beliefs, attitudes or values of others, the measures used to assess these opinions need to be thoughtfully developed.

This investigation represents a fresh approach to measuring important affective variables and using the results as input for potential change. The study was developed on solid theoretical ground and proposed a modification of Fishbein's model that improves scoring procedures for the belief and evaluation dimensions. Additional research should be attempted using this promising approach.